

RECEIVED 23 JUL 2004

**AN APPARATUS AND METHOD FOR VACUUM-BASED NANOMECHANICAL  
ENERGY, FORCE, AND MASS SENSORS**

5 *Related Applications*

The present application is related to U.S. Provisional Patent Applications serial no. 60/379,536 filed on May 7, 2002; serial no. 60/379,542, filed on May 7, 2002; serial no. 60/379,544, filed on May 7, 2002; serial no. 60/379,535, filed on May 7, 2002; serial no. 60/379,546, filed on May 7, 2002; serial no. 60/379,644, filed on May 7, 2002; serial  
10 no. 60/379,713, filed on May 7, 2002; serial no. 60/379,709, filed on May 7, 2002; serial no. 60/379,685, filed on May 7, 2002; serial no. 60/379,550, filed on May 7, 2002; serial no. 60/379,551, filed on May 7, 2002; serial no. 60/419,617, filed on Oct. 17, 2002, which are incorporated herein by reference and to which priority is claimed pursuant to 35 USC 119.

15

*Incorporation of Copending Applications*

It is to be expressly understood that the present application incorporates by reference simultaneously filed applications serial no. (PAU.35), entitled, "A Method And Apparatus For Providing Signal Analysis Of A Bionems Resonator"; and serial no.  
20 (PAU.36), entitled "Dynamics Bionems Sensors And Arrays Of BIONEMS Sensor Immersed In Fluids" as if set out in their entirety. Further, the present application incorporates by reference U.S. Patent Application serial no 10/138,538, filed on May 3, 2002 entitled, "An Apparatus and Method for Ultrasensitive Nanoelectrochemical Mass Detection"; and U.S. Patent Application serial no. 09/927,779, filed on Aug. 9, 2001,  
25 entitled, "Active NEMS Arrays for Biochemical Analyses" as if set out in their entirety.

## Background of the Invention

### 1. *Field of the Invention*

The invention relates to the field of vacuum-based nanomechanical detectors which convert some aspect or attribute of energy, force, and mass into an electrical response.

### 2. *Description of the Prior Art*

Thin, suspended two-dimensional electron gas heterostructures have been recently perfected, and have subsequently been employed for nanoscale conducting devices as described in **Blick et.al.**, Phys. Rev. B 62. In **Beck et.al.**, Appl. Phys. Lett. 68, 3763 (1996) and Appl. Phys. Lett. 73, 1149 (1998), a stress sensing field effect transistor was integrated into a cantilever and was used as deflection readout. The FET employed had transconductance of about 1000  $\mu$ S and a small signal drain-source resistance of about 10M $\Omega$ , and its strain sensitivity was presumed to arise from the piezoelectric effect.

The sensitive detection of motion in resonant mechanical systems invariably relies on at least one of the following: efficient transduction of the motion to an electrical signal, and the use of a low noise electrical readout circuit. In general, for micron-scale structures with extremely high aspect ratios operating in vacuum, transduction is sufficiently responsive to enable the detection of the structure's thermomechanical fluctuations. However, as the device size is reduced to the nanometer-scale, it becomes increasingly difficult to maintain the necessary aspect ratios for the responsive transduction required to attain the fundamental sensitivity limits of thermomechanical fluctuations or quantum zero-point motion.

The detection sensitivity in a nanoelectromechanical device, then, is in general limited by noise at the input of the linear electrical amplifier in the readout circuit, rather than by intrinsic fluctuations. In order to circumvent this limitation, it is necessary to amplify the signal by a nonlinear amplifier prior to its transmission to the linear electrical amplifier. Fortunately, one of the primary attributes of nanoelectromechanical systems (NEMS) is easily accessible nonlinearity.

Mechanical parametric amplification has been demonstrated in a few microfabricated systems in the past decade. In all these systems, amplification of the motion of the resonator was obtained by modulating the spring constant of the resonator at twice its natural frequency. The distinguishing features of these systems are their bandwidth, dynamic range, and the nature of the modulation in spring constant. Rugar and Grutter were the first to demonstrate mechanical parametric amplification in a microfabricated device. In their device, the electrical component of a silicon cantilever's spring constant was modulated by forming a capacitor between the cantilever and a baseplate, and varying the voltage between electrodes on the two surfaces. The bandwidth of their device was  $\omega_0/4Q = 5.3$  Hz and their detection sensitivity was sufficient to achieve the first demonstration of thermomechanical noise squeezing. Dana et al. observed parametric amplification in a partially metallized gallium arsenide cantilever bent by residual stress due to thermal mismatch between the metal and the gallium arsenide. Modulation of the spring constant was achieved by the superposition of a large pump drive on top of the small mechanical signal to be amplified, in order to access second-order geometric nonlinearity resulting from the curved geometry. The bandwidth in this experiment was again on the order of 6 Hz. Carr et al demonstrated parametric amplification in a surface micromachined torsional resonator operating at 500 kHz, with bandwidth on the order of 1 kHz. In this device, a

capacitor was formed between the resonator and the substrate, and the electrical component of the spring constant was again modulated by a pump signal applied across the capacitor. All these experiments showed mechanical gain from up to 20, with threshold pump voltages ranging from 200 mV to a few V.

5

#### *Balanced Electronic Displacement Detection for VHF NEMS*

The recent efforts to scale microelectromechanical systems (MEMS) down to the sub-micron domain have opened up an active research field, drawing interest from both technical and scientific communities. These nanoelectromechanical systems (NEMS) possess fundamental mechanical resonance frequencies reaching into the microwave bands and are suitable for a number of important technological applications such as ultrafast actuators, sensors, and high frequency signal processing components. Experimentally, they are expected to make possible investigations of new phonon mediated mechanical processes and of the quantum behavior of mesoscopic mechanical systems.

Among the most needed elements for developing NEMS based technologies, as well as for accessing the interesting experimental regimes they open up, are sensitive, wide-band, on-chip transduction methods sensitive to sub-nanometer displacements. While displacement detection at the scale of MEMS has been successfully realized using magnetic, electrostatic and piezoresistive transducers through electronic coupling, most of these techniques become insensitive at the sub-micron scales. Moreover, the attractive electronic two-port actuation-detection configuration of most MEMS devices becomes hard to realize at the scale of NEMS, due to the unavoidable stray couplings encountered with the reduced dimensions of NEMS.

An on-chip displacement transduction scheme that scales well into the NEMS



domain and offers direct electronic coupling to the NEMS displacement is magnetomotive detection. Magnetomotive reflection measurements on radiofrequency (RF) NEMS have found extensive use and been analyzed in detail. The operational circuit for such a measurement is shown in Fig. 19(a), with the NEMS modeled as a parallel  $RLC$  network. When driven by a source at  $\omega$ , a voltage on  $R_L$  can be detected as

$$V_0(\omega) = V_{in}(\omega) \frac{R_e + Z_m(\omega)}{R_L + 2(R_e + Z_m(\omega))} \cong V_{in}(\omega) \frac{R_e + Z_m(\omega)}{R_L + 2R_e} \quad 4.1$$

Here,  $R_e$  is the electronic DC coupling resistance to the NEMS device,  $Z_m(\omega)$  is the mechanical impedance of the resonator,  $R_L$  and  $R_s$  are the source and load impedances, respectively and the simplifying assumption  $R_L = R_s = 50 \Omega$  has been made. We have made the approximation that  $R_e \gg |Z_m(\omega)|$ , as is the case in most experimental systems. Apparently, the measured EMF due to the NEMS displacement proportional to  $Z_m(\omega)$  is embedded in a background voltage proportional to  $R_e$ . This facilitates the definition of a useful parameter, the *detection efficiency* at the mechanical resonance frequency as the ratio of the signal voltage,  $S$ , to the background,  $B$ ,

$$\frac{S}{B} = \frac{R_m}{R_e} \quad 4.2$$

The above expressions indicate some limitations of the reflective, one-port magnetomotive displacement detection. First, detection of the EMF becomes extremely challenging in interesting NEMS devices without metallization layers or having high resonance frequencies (small mechanical impedances), *i.e.* when  $R_e \gg R_m$ . Second, the voltage background in the signal prohibits the use of the full dynamic range of the detection electronics. A two-port configuration for displacement actuation and detection

might seem to remedy the above problems by improving *SIB*, but in reality, the stray electronic coupling between the ports typically dominates the measured response.

#### *Ultra High Frequency Silicon Carbide Nanomechanical Resonators*

5            Significant efforts have been made recently in the fabrication and measurement of nanomechanical resonators with fundamental resonance frequencies reaching into the UHF (ultra-high frequency) and microwave bands. Such research and development carry great importance both scientifically and technologically. In terms of fundamental science, such devices offer intriguing potential for testing quantum mechanics by  
10   observing mesoscopic mechanical motion, and for ultrasensitive measurement over the standard quantum limit. On the technological side, the nanoelectromechanical systems (NEMS), when used as high resolution sensors and actuators, or as high speed signal processing components, offer great advantage of much greater integratability over what has been implemented in today's industry.

15           Carr et al at Cornell University have recently reported successful measurement of single suspended wires with fundamental resonant frequencies up to 380 MHz. However, as indicated in their paper, "Wires with lengths below 2  $\mu\text{m}$  could not be easily detected.", which implies that 380 MHz is nearly the highest fundamental resonance frequency accessible with their technique, without major new developments  
20   to be made in the future.

#### *Frequency Tuning of MEMSINEMS Resonators by the Lorentz Force*

             High performance sensor and transducer applications of MEMS require that the device frequencies be tuned or adjusted after fabrication. Several different methods  
25   realizing tuning up to a few times the mechanical resonances have been presented for

device frequency tuning in the MEMS literature. These methods can be classified into two categories, those that alter and those that supplement the restoring forces provided by the mechanical springs. The simplest example for the former method comes from thermal cycling of a clamped beam. As the beam contracts or expands depending on the temperature change, the resonance frequency shifts due to the stress induced in the beam. The latter case has been realized by implementing electrostatic actuators in a micromechanical device that provides an electrostatic restoring force in conjunction with the mechanical spring force.

Since higher mechanical resonance frequencies in NEMS devices imply higher spring constants, force tuning by altering the mechanical restoring forces is expected to be a smaller effect in high frequency resonators. In order to assess the tuning prospects of high frequency MEMS ( $f > 1$  MHz), we have made some investigations of the dependence of the device frequencies on constant forces and temperature variations. Our measurements indicate that tuning effects indeed become obscured among other effects such as thermal frequency shifts as the device frequencies go up. Above resonance frequencies of 5 MHz, force tuning was not possible using our current techniques. In lower frequency resonators ( $1 \text{ MHz} < f < 3.5 \text{ MHz}$ ), stresses in the structures, induced during micromachining as well as in the electrical contact layers might be governing low force tuning applications. Thermal tuning also depends very strongly on the device frequency, with the largest spring constant devices showing the least tuning.

*Ultimate Limits Of Displacement Detection With Flexural And Torsional Resonators Using Magnetomotive Transduction*

Micromechanical devices have been incorporated into a wide variety of electronic

devices operating at frequencies of 1-100 kHz. Consequently, there exists a host of well-established motion detection techniques suitable for this frequency range. Since nanomechanical devices operating above 100MHz are expected to play an important role in RF signal processing, it is necessary to thoroughly characterize these techniques in this frequency range. The utility of a particular detection technique relies on three components: (1) the efficient transduction of the motion to a measurable signal, (2) the efficient coupling of that signal to the measurement apparatus, and (3) the availability of a low noise detector. What is needed is some way to quantify the performance of the magnetomotive detection technique in the context of micromechanical resonators.

#### *NEMS Array Scalar Analyzers/ Correlators*

The concept behind a mechanical array spectrum analyzer is many decades old. In one well known embodiment the analyzer functions through resonant reeds (cantilevers) that are vibrationally or electrostatically driven by an applied time-varying waveform. If the signal contains spectral weight within the band over which a given element can resonantly respond, motion of that specific element results and the amplitude of motion is proportional to the spectral weight in that band. A common application for these devices was as a tachometer, e.g. for rotary machinery, in which case an AC voltage derived from a shaft encoder is used to drive the reed array electrostatically.

Miniature suspended devices can form the basis for extremely sensitive bolometric detectors due to their miniscule heat capacities, very small thermal conductances, and the extremely fast thermal response times that result from these twin attributes. The prior art has used these attributes to demonstrate a microscale

MEMS array IR imager. There the elements were read out mechanically; upon absorption of IR radiation, an overlayer provided differential thermal expansion compared to the underlying cantilever devices. The strain induced bending was then detected by a separate optical displacement readout scheme. Other work in this area

5 has been based on thermoelectric voltages induced between different materials patterned atop suspended microscale devices. In this case, although the readout is electrical, the enhanced functionality is still derived from the small (micro scale) nature of the isolated sensing elements.

What is needed is a reinvigoration of such analyzers through access to

10 mechanical response from UHF to microwave frequencies using NEMS technologies, to give the prospect of ultralow operating power levels and the monolithic, ultracompact form.

### **Brief Summary of the Invention**

15 Nanoelectromechanical systems, or *NEMS*, are mechanical devices scaled to submicron dimensions. In this size regime, it is possible to attain extremely high fundamental frequencies while simultaneously preserving very high mechanical responsivity (small force constants) and reasonable quality factors (*Q*) for the resonant mechanical response. This powerful combination of attributes translates directly into

20 optimal characteristics for mechanical sensing, *e.g.*

- a) high energy, force, and mass sensitivity
- b) operability at ultralow power
- c) the ability to induce usable nonlinearity with quite modest control forces.

NEMS thus engender electromechanical device applications requiring fast response times; operating frequencies comparable to most of today's purely electronic devices are attainable.

Multiterminal electromechanical devices are possible, *i.e.* devices that

5 incorporate two-, three-, four-ports. In these, separate electromechanical transducers can provide both input stimuli, *i.e.* signal forces, and readout of the mechanical response, *i.e.* output displacement. Hereafter these are termed *actuators* and (displacement) *transducers*, respectively. Through additional *control transducers*, electrical signals — either quasi-static or time-varying — can be applied and converted

10 into quasi-static or time-varying forces that excite or perturb the properties of the mechanical element in a controlled and useful manner. Utilizing different physical processes of electromechanical transduction and actuation allow highly independent interaction between these ports, in effect enabling "orthogonality" between the input, output and possibly multiple control ports. In other words, each port can strongly

15 interact with the mechanical element, while maintaining relatively weak direct couplings to each other. For time-varying stimuli when frequency conversion is the goal, this orthogonality can be provided by a tuned or narrowband transducer response to (frequency-) select input and output signals from control signals, *e.g.* pump signals.

#### 20 *Transduction between signal domain and displacement.*

An output signal in displacement domain can be a static shift, resonant response, modulation of steady-state induced vibration amplitude, modulation of the harmonic content of steady-state induced vibration, or modification of noise spectrum, etc. The following table represents the range of models for transduction:

<b>Input Signal Domain</b>	<b>Sensing Modality (relevant responsivity)</b>
Energy	energy loss (damping, Q factor)
Thermal parameters	increase in thermomechanical noise
Force	static displacement (compliance) and/or resonant displacement (dynamical compliance)
Mass changes	frequency shift (mass responsivity)

### *Elements of Nanomechanical Sensors*

#### *Compliant Elements*

The compliant elements are the mechanical structures scaling down to submicron size which move or are displaced. Due to their extremely small size, they act as efficient probe to the microscopic world. These structures are usually made of semiconductor materials. For example, in this invention, we have used GaAs, Si, SiC, and GaAs/AlGaAs heterostructures. Sometimes, pure metal or metal alloy can be used. The selection of materials depends largely on their electrical, chemical and mechanical properties. Sensor geometry is an important factor in the designing. Finite element simulation is useful in the estimate of the resonant frequency, spring constant, force/mass sensitivity.

#### *Transducers*

The structure which produces a piezoelectric, piezoresistive, magnetomagnetic or other transformation from the input signal domain to the sensing modality comprises

the transducer. Typically, this is a compositional structural layer or a current path and source for generating a Lorentz-force-derived emf.

### *Actuators*

5           The structure which produces the mechanical movement of the NEMS device is the actuator, which may be an external current and magnetic field combination for the driving Lorentz force in a magnetomotive transducer, a current generating a dipole field on an adjacent electrode, or even stochastic thermal fluctuations of an ambient fluid.

### 10   *Nanomechanical Sensor Systems*

Sensor systems comprise simple one element systems, or more complex compound-element designs to achieve specific functionality. The sensed electrical signal generated in or the changed electrical parameter of the transducer may be sensed in a bridge, one port, two port or other multiple port combination.

15           “NEMS” in this specification is used to mean devices with at least one dimension which is equal to or smaller than one micron. It does not exclude the possibility that the “NEMS” device may have one or more other dimensions larger than one micron. Furthermore, as can be understood there is often no sharp line of distinction between the characterization of a device at or below one micron in size and one which is above  
20   one micron. The more meaningful significance to the term, “NEMS” that the device in question shares some characteristic with similar devices scaled to submicron sizes or which is unique to submicron devices or operation.

The invention is directed to an apparatus and method which produces a high resolution displacement readout that is based upon our ability to achieve very high  
25   mobility suspended quantum wires. Two-terminal sensor impedances as low as 5k $\Omega$ .



Molecular beam epitaxial (MBE) grown materials are directly patterned and in-plane gates (IPG) are used to excite the vibration. No metallization is needed. Hence high Q values can be obtained.

5 The mechanical parametric amplifier described is a practical solution to the problem of detection sensitivity, as it utilizes the geometric nonlinearity inherent in NEMS.

The invention is more specifically defined as a monolithically fabricated apparatus comprising a doubly clamped, suspended beam with a submicron width having an asymmetrically positioned, mechanical-to-electrical transducing layer  
10 fabricated within or on the beam. At least one side drive gate is provided proximate to the beam within a submicron distance.

The asymmetrically positioned, mechanical-to-electrical transducing layer comprises an asymmetrically positioned piezoelectric layer within the beam. The beam is fabricated from a 2 DEG heterostructure.

15 In one embodiment the beam is provided with electrical contacts and forms a two-terminal circuit with an output terminal, and further comprises an inductor in parallel circuit with the beam and a blocking capacitor coupled to the output terminal of the beam. A low noise cryogenic amplifier is coupled to the blocking capacitor.

The gate is provided with a gate dipole charge separation and the beam is  
20 provided with a beam dipole charge separation, so that the beam and gate interacting through the dipole-to-dipole interaction. The side gate includes a 2 DEG layer.

In the illustrated embodiment the beam and side gate comprise a chip and further comprise a substrate on which the chip is disposed, the substrate having an electrode formed thereon, where the gate being provided with a gate dipole charge  
25 separation between the electrode of the substrate and the gate. The beam is provided

with a beam dipole charge separation, the beam and gate interacting through the dipole-to-dipole interaction.

In one embodiment the beam and gate are fabricated from an asymmetric heterostructure stack comprising a 2 DEG GaAs piezoelectric layer, two sandwiching  
5 AlGaAs spacer layers on each side of the GaAs layer, a first and second AlGaAs: Si donor layer above and below the AlGaAs spacer layers respectively, two GaAs cap layers above and below the AlGaAs: Si donor layers respectively. Each of the layers below the 2 DEG GaAs piezoelectric layer is thicker than the corresponding layer above the 2 DEG GaAs piezoelectric layer. An  $\text{Al}_x\text{Ga}_{1-x}\text{As}$  sacrificial layer is disposed under  
10 the stack and a substrate disposed under the  $\text{Al}_x\text{Ga}_{1-x}\text{As}$  sacrificial layer, where  $0 < x < 1$ .

The apparatus may further comprise two gates, each disposed within a submicron distance of the beam and each provided with a gate dipole charge separation.

15 The apparatus further comprises a source of sensing current supplied to the beam and an amplifier in circuit with the beam to generate an output signal. In the illustrated embodiment the amplifier is cryogenic.

The source of sensing current supplies a DC and AC sensing current to the beam.

20 In one embodiment transducing layer of the beam is piezoelectric which is used to induce oscillation of the beam, and is also piezoresistive which is used to sense oscillation of the beam.

The invention is still further defined as an improvement in a method of forming a suspended NEMS beam including a two-dimensional-electron-gas layer comprising the  
25 steps of providing a heterostructure stack including a 2 DEG layer disposed on a

sacrificial layer; selectively disposing a mask on the stack to define a pattern for the NEMS beam; dry etching away exposed portions stack the using a  $\text{Cl}_2/\text{He}$  plasma etch to define the NEMS beam without substantially altering the electrical characteristics of the 2 DEG layer; and etching the sacrificial layer away to release the NEMS beam.

5           The step of dry etching away exposed portions stack the using a  $\text{Cl}_2/\text{He}$  plasma etch comprises supplying  $\text{Cl}_2$  and He gas at a flow rate ratio of 1:9 respectively into an ECR plasma chamber.

          The step of supplying  $\text{Cl}_2$  and He gas into the ECR plasma chamber further comprises maintaining the stack at or less than 150V self-bias with 20W constant RF  
10   power and ionizing the  $\text{Cl}_2$  and He gas with approximately 300W microwave power or more.

          The invention is also a NEMS parametric amplifier comprising: a suspended oscillating submicron signal beam defined in a plane and having a flexural spring constant for in-plane motion and being driven at  $\omega$  at or near the frequency of  
15   mechanical resonance of the signal beam; a pair of pump beams coupled to the signal beam and being driven at or near  $2\omega$ ; a source of magnetic field applying a field with at least a component perpendicular to the signal beam and pair of pump beams; and a source of alternating current coupled in circuit with the pump beams to apply a current through the pump beams in the presence of the magnetic field to generate a modulated  
20   Lorentz force on the pump beams to apply in turn a force oscillating of compression and tension to the signal beam to perturb the flexural spring constant for in-plane motion of the signal beam. An amplifier may be coupled to the beam.

          The pump beams and signal beam collectively form an H-shaped structure in the plane, the signal beam forming the middle portion of the H-shaped structure. The pump  
25   beams are tuned to resonate at  $2\omega$ .

The invention is also a method of operating the NEMS parametric amplifier described above.

The invention is also a submicron cantilever characterized by a submicron displacement comprising a NEMS cantilever having a restriction portion; a

- 5 piezoresistive strain transducer epilayer coupled to the cantilever; where G is the gauge factor of the apparatus given by

$$G = \frac{3\beta\pi_L K(2l - l_1)}{2bt^2} R_T$$

- where the parameter  $\pi_L$  is the piezoresistive coefficient of the piezoresistive transducer material, the factor  $\beta$  accounts for the decrease in G due to the finite  
10 thickness of the conducting layer, K is the spring constant of the cantilever, l the overlength of the cantilever,  $l_1$  the length of the restriction portion, b the thickness of the restriction portion, t the thickness of the thickness of the restriction portion, and  $R_T$  is two-terminal resistance of the transducer.

- Near resonance, the force spectral density of thermomechanical fluctuations is  
15 given by

$$S_F^\gamma = 4k_B T \gamma = 4Kk_B T l / (2\pi Q f_0)$$

where  $k_B$  is the Boltzman constant, T is the temperature,  $\gamma$  is the damping coefficient,  $f_0$  is the resonance frequency and  $Q = mf_0/\gamma$  is the quality factor, m is the mass of the cantilever.

- 20 Near resonance, the voltage spectral density for the thermomechanical fluctuations is given by

$$S_V^\gamma = \frac{S_F^\gamma G^2 l^2}{16\pi^2 m^2 f_0^2 [4(f - f_0)^2 + f_0^2 / Q]}$$

where f is the frequency of oscillation of the cantilever.

The invention is a method for scaling and determining carrier distribution in NEMS devices having a doped layer with different doping concentration and different thicknesses disposed on an intrinsic layer comprising the steps of: providing the doped layer with a predetermined thickness; providing a doping concentration in the doped layer; adjusting the Fermi level until charge neutrality is obtained by satisfying the condition

$$\int_0^L (\rho(x)/e + N_A^-(x)) dx = 0$$

where

$$N_A^-(x) = \frac{\# \text{ dopants}}{\frac{1}{2} e^{-\beta(E_A - (E_F - E_V))}}$$

10 is the density of ionized acceptor sites, where  $p$  is volume density of carriers given by Fermi statistics,  $\rho(x) = e(p(x) - n(x))$  and positive and negative carrier densities are

$$p(x) = 1.04 \times 10^{25} e^{-\beta(E_F - E_V)} / m^3$$

$$n(x) = 2.8 \times 10^{25} e^{-\beta(E_C - E_F)} / m^3$$

where  $\beta$  is  $1/kT$ ,  $E_F$  is the Fermi energy,  $E_V$  is the energy of the valence band energy, and  $E_C$  is the energy of conduction band; determining the bending of the valence band according to the equation

$$\frac{d^2 E_v}{dz^2} = \frac{e \rho(x)}{\epsilon}$$

where  $E_v$  is the energy of the valence band,  $\epsilon$  is the dielectric constant,  $e$  is the charge of the electron, subject to the boundary condition:

$$\left. \frac{d^2 E_v}{dz^2} \right|_{z=0} = \frac{e \sigma}{\epsilon}$$

where  $\sigma$  is the empirical surface carrier density; and iteratively repeating the foregoing steps of adjusting and determining until convergence is attained for a carrier density,  $\rho$ .

The invention is also a bridge circuit comprising: a source of excitation signal; a  
5 power splitter coupled to the source to generate two out-of-phase components of the excitation signal; a first actuation port coupled to the power splitter; a second actuation port coupled to the power splitter; a first circuit arm coupled to the first actuation port including a first NEMS resonating beam having an transduced electrical output; a  
10 second circuit arm coupled to the second actuation port including a second NEMS resonating beam having an transduced electrical output, the first and second beams being matched to each other; and a detection port coupled to the DC coupling resistance,  $R_e$  and to the NEMS resonating beam.

The bridge further comprises a variable attenuator and a phase shifter coupled in circuit in opposing ones of the first and second circuit arms. The attenuator balances  
15 out impedance mismatch between the first and second circuit arms more precisely than without the inclusion of the attenuator, while the phase shifter compensates for the phase imbalance created by the circuit inclusion of the attenuator.

The NEMS resonating beam includes a surface adapted to adsorb a test material, performance of the NEMS resonating beam being affected by the test material  
20 and being measured by the bridge.

The bridge further comprises an amplifier and an output impedance mismatch circuit coupling the detection port to the amplifier. The first and second NEMS resonating beams are magnetomotive NEMS resonating beams and have no metallization.

The invention is still further a method of balancing the output of two NEMS devices in a bridge circuit as described above.

The invention is defined as an apparatus comprising a driving source; a power splitter coupled to the source for generating drive signals of opposing phases; a first  
5 magnetomotive NEMS resonating beam coupled to one phase of the drive signal generated by the power splitter; a second magnetomotive NEMS resonating beam coupled to the other opposing phase of the drive signal generated by the power splitter; a terminal electrical coupled to the two magnetomotive NEMS resonating beams; an amplifier coupled to the terminal; and means coupled to the amplifier, the means for  
10 measuring the frequency dependence of the forward transmission coefficient  $S_{21}$  of the apparatus.

The first and second magnetomotive NEMS resonating beams are comprised of SiC and which vibrate in an in-plane resonance and in an out-of-plane resonance. An adsorbing surface is disposed on one of the NEMS resonating beams, and adsorption  
15 of an adsorbate on the adsorbing surface is measured by the means for measuring.

The invention is a method comprising the steps of providing an excitation driving signal; splitting the excitation driving signal into two out-of-phase components; providing one of the out-of-phase components to a first NEMS resonating beam having a first transduced electrical output; providing the other one of the out-of-phase components to  
20 a second NEMS resonating beam having a second transduced electrical output, the first and second beams being matched to each other; vibrating the first and second NEMS resonating beams; summing the first and second transduced electrical outputs together to generated a balanced detected output signal; amplifying the balanced detected output signal in an amplifier; and measuring the frequency dependence of the  
25 forward transmission coefficient  $S_{21}$ .

The step of vibrating the first and second magnetomotive NEMS resonating beams comprises vibrating the beams at an in-plane resonance and/or at an out-of-plane resonance.

The invention is yet further defined as an improvement in a magnetomagnetically driven submicron NEMS resonating beam comprising a submicron SiC NEMS beam having a surface and an axial length  $L$ , width  $W$ , Young's modulus  $E$ , mass density  $\rho$ , and displacement amplitude  $A$ ; a source of a magnetic field,  $B$ ; an electrode means disposed on the surface of the beam for conducting current along at least a portion of the axial length of the beam; a source of alternating current coupled to a first end of the electrode means to magnetomotively drive the SiC NEMS beam to a resonant

frequency  $f_0 = \sqrt{\frac{E}{\rho}} \frac{W}{L^2}$ ; and a detector coupled to a second end of the electrode

means to detect a generated  $V_{emf}$  from the SiC NEMS beam of

$$V_{emf} \propto B A \sqrt{\frac{E}{\rho}} \frac{W}{L}.$$

The electrode means comprises a single electrode coupled to the source of alternating current for driving the beam in the magnetic field and is coupled to the detector for sensing the EMF generated in the electrode by motion of the beam.

The electrode means comprises a first electrode coupled to the source of alternating current for driving the beam in the magnetic field and a second electrode coupled to the detector for sensing the EMF generated in the electrode by motion of the beam.

The SiC NEMS beam has dimensions and parameters providing a fundamental resonance frequencies in the UHF range and higher and in particular in the microwave L band.



The invention is a method of tuning a submicron NEMS device having an out-of-plane resonance comprising providing a magnetic field in which the NEMS device is positioned; supplying an AC current to the NEMS device to oscillate the NEMS device in the magnetic field at a resonant frequency; supplying a DC current to the NEMS  
 5 device to tune the out-of-plane resonant frequency of the NEMS device with a constant Lorentz force.

The step of supplying a DC current to the NEMS device comprises supplying a DC current to the metallization.

The NEMS device also has an in-plane resonance and the method further  
 10 comprises the step of varying the temperature of the NEMS device to tune both the out-of-plane and in-plane resonance of the NEMS device.

The invention is also a tunable submicron NEMS device having an out-of-plane resonance which is tuned by the above method. The NEMS device comprises a semiconductor-metal bilayer formed of a single crystalline highly doped semiconductor  
 15 and the metallization disposed thereon is a polycrystalline metal to reduce stresses in the semiconductor-metal bilayer.

The invention is characterized as an improvement in a resonating submicron one-port NEMS device comprising a resonating beam having a width  $w$ , a thickness  $t$ , a length  $L$ , a detector load resistance  $R_L$ , an equivalent mechanical impedance  $R_m$ ,  
 20 operating a frequency corresponding to the wavelength  $\lambda$  with an electrode on the beam with a conductivity of  $\sigma$  such that the insertion loss  $\epsilon$  defined as:

$$\epsilon_1 = \frac{\alpha^2}{(1+\alpha)(1+\alpha + \frac{R_m \lambda \sigma t w}{L})} \text{ where } \alpha = \frac{\lambda \sigma R_L t w}{L}$$

is minimized or near unity.

The invention is an improvement in a resonating submicron two-port NEMS device comprising a resonating beam having a width  $w$ , a thickness  $t$ , a length  $L$ , a detector load resistance  $R_L$ , an equivalent mechanical impedance  $R_m$ , operating a frequency corresponding to the wavelength  $\lambda$  with an electrode on the beam with a conductivity of  $\sigma$  such that the insertion loss  $\epsilon$  defined as:

$$\epsilon_2 = \frac{1}{2} \alpha^{\frac{1}{2}} \left( \frac{1-\alpha}{1-0.75\alpha} \right)^{\frac{1}{2}} \text{ where } \alpha = \frac{\lambda \sigma R_L t w}{L}$$

is minimized or near unity.

The invention is an improvement in a two-port, straight, doubly clamped NEMS magnetomotive beam coupled to an amplifier with a load resistance  $R_L$ , the NEMS beam having a length  $L$ , a thickness  $t$ , a width  $w$ , Young's modulus  $E$ , mass density  $\rho$ , in a magnetic field  $B$ , with a conductivity  $\sigma$  of its metallization, a temperature  $T$ , a driving signal wavelength of  $\lambda$ , a resonant frequency of  $f_0$ , an amplifier spectral power density  $S_v^a$ , chosen so that the spectral displacement sensitivity  $S_{x(2)}^m$  is equal to or greater than the spectral displacement density corresponding to thermal fluctuations of the NEMS beam, which spectral displacement sensitivity  $S_{x(2)}^m$  is defined as

$$\sqrt{S_{x(2)}^m} = \frac{1.68}{\sigma^{\frac{1}{2}} \lambda^{\frac{1}{2}} B} \left( \frac{\rho}{E} \right)^{\frac{1}{8}} f_0^{-\frac{3}{4}} t^{-\frac{3}{4}} w^{-\frac{1}{2}} \left[ k_B T + \frac{S_v^a}{R_L} \left( \frac{1-0.75\alpha}{1-\alpha} \right) \right]^{\frac{1}{2}}$$

where  $k_B$  is the Boltzman constant and  $\alpha = 0.99 R_L \sigma \lambda \left( \frac{\rho}{E} \right)^{\frac{1}{4}} f_0^{\frac{1}{2}} t^{\frac{1}{2}} w$ .

The invention is a method for fabrication of a NEMS beam from a Si membrane comprising the steps of : providing a Si substrate; disposing a  $\text{SiO}_2$  layer on the Si

substrate; disposing a Si epilayer on the SiO<sub>2</sub> layer; selectively anisotropically etching away a portion of the Si substrate down to the SiO<sub>2</sub> layer used as a stop layer; selectively etching away a portion of the SiO<sub>2</sub> layer to expose a suspended Si epilayer membrane; and forming the NEMS beam in the suspended Si epilayer membrane, 5 whereby capillary distortion is avoided and electron beam resolution is achieved without proximate scattering from a substrate.

The invention is a method for fabrication of a NEMS beam from a GaAs membrane comprising the steps of providing a GaAs substrate; disposing an AlGaAs layer on the GaAs substrate; disposing a GaAs epilayer on the AlGaAs layer; 10 selectively anisotropically etching away a portion of the GaAs substrate down to the AlGaAs layer used as a stop layer; selectively etching away a portion of the AlGaAs layer to expose a suspended GaAs epilayer membrane; and forming the NEMS beam in the suspended GaAs epilayer membrane.

The step of selectively anisotropically etching away a portion of the GaAs 15 substrate down to the AlGaAs layer used as a stop layer comprises etching with a NH<sub>4</sub>OH or citric acid solution. The step of etching with a NH<sub>4</sub>OH solution comprises etching with a solution comprised of NH<sub>4</sub>OH and H<sub>2</sub>O<sub>2</sub> in the volume ratio of approximately 1:30, freshly mixed prior to etching.

The step of etching with a citric acid solution comprises etching with a room 20 temperature bath comprised of citric acid monohydrate mixed and completely dissolved in a 1:1 mixture with deionized water by weight, then mixing this 1:1 mixture in a 3:1 volume ratio with H<sub>2</sub>O<sub>2</sub> to provide the bath.

The invention is a NEMS array analyzer comprising two opposing parallel substrates; a plurality of piezoresistive NEMS cantilevers extending from one of the 25 substrates, each of the NEMS cantilevers having a different resonant frequency so that

the corresponding plurality of resonant frequencies covers a selected spectral range; and a plurality of drive/sense elements extending from the other one of the substrates, each of the drive/sense elements primarily coupled with one of the plurality of piezoresistive NEMS cantilevers.

5           *The invention is a NEMS array analyzer comprising a frame; a plurality of NEMS structures forming an interacting array to form an optical diffraction grating; means for driving the plurality of NEMS structures in response to an input signal; and light source for illuminating the plurality of NEMS structures; and detector means for detecting*  
10 *diffracted light from the plurality of NEMS structures acting collectively as a time-varying diffraction grating.*

          The invention is a NEMS electronic chemical sensing array comprising a plurality of strain-sensing NEMS cantilevers, each having an overlayer disposed thereon which is responsive to a corresponding analyte, the response of the overlayer imposing a strain on the corresponding cantilever; and means for detecting the strain of each of the  
15 plurality of strain-sensing NEMS cantilevers. The response of the overlay comprises expansive or contractile volume changes of the overlay causing a strain to be imposed on the corresponding cantilever to cause it to bend, and where the means for detecting comprises an optical detector array for determining the amount of bending of each cantilever. The response of the overlay comprises a mass loading resulting in a change  
20 in total inertial mass of each corresponding cantilever and where the means for detecting comprises means for detecting changes in resonant frequency shifts for each cantilever.

          The invention is a NEMS infrared sensing array comprising: two opposing parallel substrates; a plurality of identically sized piezoresistive NEMS cantilevers  
25 extending from one of the substrates, each of the cantilevers being provided with a

corresponding IR absorber responsive to a different IR frequency and inducing a corresponding differential thermal expansion of each cantilever depending on the amount of IR absorbed by each IR absorber; and a plurality of drive/sense elements extending from the other one of the substrates, each of the drive/sense elements  
5 primarily coupled with one of the plurality of piezoresistive NEMS cantilevers.

While the apparatus and method has or will be described for the sake of grammatical fluidity with functional explanations, it is to be expressly understood that the claims, unless expressly formulated under 35 USC 112, are not to be construed as necessarily limited in any way by the construction of "means" or "steps" limitations, but  
10 are to be accorded the full scope of the meaning and equivalents of the definition provided by the claims under the judicial doctrine of equivalents, and in the case where the claims are expressly formulated under 35 USC 112 are to be accorded full statutory equivalents under 35 USC 112. The invention can be better visualized by turning now to the following drawings wherein like elements are referenced by like numerals.

15

### **Brief Description of the Drawings**

Fig. 1a is a graph of the energy band level in a heterostructure as shown in Fig. 1b at different points in the thickness,  $t$ .

Fig. 1b is a side cross-sectional diagram illustrating the stack in which the NEMS  
20 device of the invention is built.

Fig. 2 is a cross-sectional schematic of the dipolar actuation mechanism of the invention, showing dipole formation on the beam between  $p_1$  of the beam and  $dp_2$  and on the driving gate.

Fig. 3(a) scanning electron microscope image of a doubly clamped beam used in  
25 the invention. The in-plane gates are formed by the 2DEG.

Fig. 3b is a schematic of the measurement setup.

Fig. 3c is a simplified side cross-sectional view of an ECR chamber used in the plasma etching step of the invention.

Fig. 3d(i) – (v) is a series of perspective views illustrating the steps of fabricating the 2DEG used in the heterostructure of Fig. 1b.

5 Fig. 4a is a graph of the voltage drop across the beam verses frequency as it is driven to its lowest mechanical resonance with increasing drive amplitudes. The DC bias current is fixed at 5 $\mu$ A. In the inset the peak value of amplitude response is shown as a function of driving amplitude in the linear regime.

10 Fig. 4b is a graph the magnitude response curve verses frequency at various DC bias currents. In the inset the signal amplitude at resonance with a sensing current increase form -26  $\mu$ A to 26  $\mu$ A.

Fig. 5 is a graph of the magnitude response curve verses frequency at various temperatures.

15 Fig. 6 is a microphotograph of the mechanical preamplifier fabricated by surface nanomachining of a 200 nm thick layer of silicon carbide on silicon. The metallic electrodes are patterned from a 50 nm thick layer of Au.

20 Fig. 7 is a diagram which illustrates the operational principals for the all-mechanical parametric amplifier. The signal electrode is used for excitation and detection of the signal beam, while the pump electrode modulates its flexural spring constant.

Fig. 8 is a circuit schematic of the circuit employed for gain measurements for the parametric amplifier in the illustrated embodiment.

25 Fig. 9 is a graph of the frequency shift  $\Delta f/f$  as a function of transverse DC force applied to the pump beams. The force is effectively a compressive (positive) or tensile (negative) force on the signal beam. The linear component of frequency shift results

from this force, while the quadratic component results from ohmic heating due to current in the pump beams.

Fig. 10 is a diagram of a finite element simulation of the parametric amplifier under a static load of 1 nN applied to the pump beams arising from the compressive or  
5 tensile force on the signal beam described in Fig. 9. The compression of the signal beam is 0.235 times what would be expected if the pump beams were not present and the load were applied directly to the ends of the signal beam.

Fig. 11 is a graph showing the dependence of the gain on the phase difference between signal and pump excitation. Depending on the phase, the signal is either  
10 amplified or de-amplified. As expected, the magnitude of both amplification and de-amplification increases for stronger magnetic fields.

Fig. 12 is a graph of the response of the signal beam to excitation at frequencies off-resonance, with the pump beams driven at twice the resonance frequency. The plot shows the strength of the sideband at  $\omega$ . The device bandwidth is reduced dramatically  
15 for pump excitations near threshold.

Fig. 13 is a graph of the amplification of thermomechanical noise. At the pump voltage of 8.2 mV, the  $\Phi=0$  gain is 39, and the quality factor of the resonance is increased from 10600 to 180000.

Fig. 14 are phasor plots of the output noise for the parametric mechanical  
20 amplifier. The top left plot shows the lock-in amplifier measurement of the signal beam with no excitation, and no pump. This displays the phase-independent input noise of the amplifier. The top right plot shows the measurement of the signal beam with no excitation and 5 mV pump voltage. The fluctuations are still dominated by the electrical amplifier. The bottom left plot shows the measurement of the signal beam with no  
25 excitation and pump voltage of 8.1 mV. Thermomechanical fluctuations are amplified

beyond the amplifier input noise in one quadrature. In the other quadrature, the effect of the pump is not seen.

Fig. 15 is a graph which provides a comparison of gain with the noise level in each quadrature, normalized to the values with pump off. The effect of the pump is to  
5 increase the signal-to-noise ratio, especially with respect to the  $\Phi=\pi/2$  quadrature.

Fig. 16 is a graph which shows the dependence of the gain on the voltage applied to the pump. At low pump amplitudes, the gain is independent of the excitation of the signal beam. At high pump voltages, the gain begins to saturate when the rms amplitude of motion reaches 360 pm.

10 Fig. 17 is a graph of the carrier distribution for a sample of 130 nm thickness in which the dopant layer is 30nm thick and the dopant concentration is  $4 \times 10^{25} \text{ m}^{-3}$ .

Fig. 18 is a graph of the carrier distribution for a sample of 30 nm thickness in which the dopant layer is 7nm thick and the dopant concentration is  $4 \times 10^{25} \text{ m}^{-3}$ .

Figs. 19a, 19b, 19c and 19d are directed to magnetomotive reflection and bridge  
15 measurements. Fig. 19a is a schematic diagram illustrating the magnetomotive reflection and Fig. 19b is a schematic diagram illustrating bridge measurements. Fig. 19c is a scanning electron microscope (SEM) micrograph of a representative bridge device of Fig. 19b. Fig. 19d is a schematic illustration of the reflection and bridge  
20 arrangements, showing perspective views of the single and balanced beam configurations respectively.

Fig. 20a is the graph of a doubly-clamped, B-doped Si beam resonating at 25.598 MHz with a Q about  $3 \times 10^4$  measured in reflection in the upper curves and in bridge configurations for magnetic field strengths of  $B=0,2,4,6 \text{ T}$  in the lower curves. Fig. 20b  
25 is a graph of the amplitude of the broadband transfer functions for both reflection and bridge configurations.



Fig. 21 is a graph of the amplitude of transmission coefficient ( $S_{21}$ ) measured from SiC beams in the bridge configuration for different magnetic field strengths of  $B=2,4,6,8$  T.

5 Figs. 22a – 22d are SEM micrographs of one embodiment of the device. Fig. 22a is a top plan view. Fig. 22b is a plan side view. Fig. 22c is an enlarged top plan view of one of the beams. Fig. 22d is an enlarged side plan view of one of the beams showing clear suspension of the mechanical structure.

Fig. 23 is a schematic drawing of measurement setup.

10 Fig. 24 is a three-dimensional graph of the frequency dependence of the forward transmission coefficient  $S_{21}$  of the network under study. The insert shows the projection of the complex function onto the  $S_{21}$  plane.

Fig. 25 is a graph of the signal amplitude referred back to the input of the pre-amplifier. This is obtained by taking modulus after subtracting the background function from the raw data, see text for the procedure of subtraction.

15 Fig. 26 is a SEM photograph showing a top plan view of the device used to illustrate high frequency tuning.

Fig. 27 is a graph of measured resonances vs. aspect ratios of Si and GaAs beams.

20 Fig. 28 is a graph of the out of plane frequency shift of a GaAs Beam with applied Lorentz force.

Fig. 29 is a graph of the frequency shift as in Fig. 28 plotted as a function of applied force.

Fig. 30 is a graph of the Lorentz Force tuning for the in plane direction.

25 Fig. 31 is a graph of the frequency shifts in Fig. 29 plotted as a function of the tuning force.

Fig. 32 is a graph of the temperature shifts of the two modes of a beam.

Fig. 33 is a graph of the temperature dependence of the resonance frequencies of three Si beams.

Fig. 34 is a graph of the temperature dependence of the resonance frequencies  
5 of four GaAs beams.

Fig. 35 is a graph of the corrected data for Fig. 29.

Fig. 36 is a schematic of an equivalent circuit for a mechanical resonance.

Fig. 37 is a schematic for an one-port drive and detection circuit.

Fig. 38 is a schematic for an equivalent circuit for one-port measurement.

10 Fig. 39 is a schematic for an equivalent circuit for a two-port detection circuit.

Fig. 40 is a simplified top views of representative designs for flexural (left) and torsional (right) resonators.

Fig. 41 is a graph of the sensitivity of the two-port magnetomotive detection technique as a function of frequency, compared to thermomechanical noise.

15 Fig. 42 is a graph of the input noise level required of a  $50\Omega$  amplifier for magnetomotive sensitivity limited by thermomechanical noise, as a function of the conductivity of the electrode.

Figs. 43a – 43d are side cross-sectional views of a method of fabricating Si membranes using bulk micromachining.

20 Figs. 44a – 44d are side cross-sectional views of a method of fabricating GaAs membranes using bulk micromachining.

Figs. 45a and 45b are SEM pictures of wells etched in GaAs with  $\text{NH}_4\text{OH}$ : Fig. 45a shows a tilted view from backside, cleaved along [011] plane, Fig. 45b shows a face-on view of [011] plane. Note the smooth, well defined sides and bottom.

25 Figs. 46a and 46b are SEM pictures of wells etched in GaAs with citric acid: Fig.

46a shows a tilted view from backside, and Fig. 46b shows a plane cleaved along the [011] plane. Note the inhomogeneity of descending walls and the roughness of floor surface. The dashed line represents the [011] cleave plane.

Fig. 47 is a simplified perspective diagram of a NEMS array based power spectrum analyzer. Elements within the array are electrostatically actuated by local stubs protruding along a common transmission line electrode. Each resonant element is separately read out piezoresistively. The element lengths are staggered, as in a vibrating reed tachometer, to provide coverage over a desired spectral range.

Fig. 48 is a diagrammatic depiction of a NEMS array spectrum analyzer based upon the collective modes arising in a coupled array. The signal is applied to the entire array, but readout is optical, and involves simultaneous resolution of the diffracted orders using a photodiode array.

Fig. 48a is an enlarged SEM photo of the array of Fig. 48.

Fig. 49 is a diagrammatic depiction of a NEMS array based electronic nose in which resonant sensors used to monitor mass loading and changes in surface strain induced by chemical or biochemical adsorbates.

Fig. 50 is a diagrammatic depiction of a NEMS array based uncooled IR imager. An array of resonant sensors is used to monitor out-of-plane flexure arising from absorption of IR energy. Local radiation induced heating of the IR absorbers results in differential thermal expansion between the absorbers and the cantilevers. The common electrostatic bias/drive connection provides a local dc electrostatic bias and a common ac drive electrode for swept frequency interrogation of the array.

Fig. 51a is a scanning electron microphotograph of a piezoelectric cantilever. The dimensions of the device are 15 $\mu$ m in length, 2 $\mu$ m in width and 130nm thickness of which the top 30nm forms the conducting layer (with a boron doping density of

$4 \times 10^{19}/\text{cm}^3$ ). For this device  $b=0.5\mu\text{m}$  and  $l_1=4\mu\text{m}$ .

Fig. 51b is a graph of cantilever displacement as a function of time, studied using an atomic force microscope tip to move the cantilever a known amount. This yields a direct measurement of  $G = dR_T / dx = 3 \times 10^7 \Omega/\text{m}$ .

5 Fig. 51c is a graph of cantilever resistance as a function of time corresponding to Fig. 51b, studied using an atomic force microscope tip to move the cantilever a known amount. This yields a direct measurement of  $G = dR_T / dx = 3 \times 10^7 \Omega/\text{m}$ .

Fig. 52 is a graph of the nanomechanical resonance peak in vacuum. The dependence of the quality factor on pressure is shown in the inset. A bias current of  
10 102 $\mu\text{A}$  was used for these measurements.

Figs. 53a and 53b is a graph of the 9K measurement of thermomechanical noise.

Figs. 54a - 54c are diagrammatic side cross-sectional views of scaled piezoresistive structures in which the scaling has been augmented with additional  
15 semiconductive layers to confine the carriers in a quantum well.

Fig. 55 is a diagrammatic side cross-sectional views of scaled piezoresistive structures in which the scaling has been augmented with a quantum well disposed on an insulator.

The invention and its various embodiments can now be better understood by  
20 turning to the following detailed description of the preferred embodiments which are presented as illustrated examples of the invention defined in the claims. It is expressly understood that the invention as defined by the claims may be broader than the illustrated embodiments described below.

## Detailed Description of the Preferred Embodiments

### *Doubly Clamped Beam*

Doubly clamped beams from GaAs/AlGaAs quantum well heterostructure containing a high-mobility two-dimensional electron gas (2DEG) is disclosed which  
5 applies an IT-drive to in-plane side gates to excite the beam's mechanical resonance through a dipole-dipole mechanism. Sensitive high frequency displacement transduction is achieved by measuring the A.C. EMF developed across the 2DEG in the presence of a constant D.C. sense current. The high mobility of the incorporated 2DEG provides low-noise, low power, and high gain microelectromechanical displacement  
10 sensing, through combined piezoelectric and piezoresistive mechanisms.

A beam 30 is formed between two gates 32 to collectively comprise a device 12 as shown in Fig. 2 and in the microphotograph of Fig. 3. The starting material was a specially designed, MBE-grown two dimensional electron gas (2DEG) heterostructure. The structural layer stack, generally denoted by reference numeral 10, from which the  
15 devices 12 of Fig. 2 are formed, comprises seven individual layers having a total thickness of 115nm as shown in Fig. 1b. The top and bottom layers 14 are thin GaAs cap layers preventing oxidation of the AlGaAs:Si donor layers 16 in between. The central 10 nm-thick GaAs layer 18 forms a quantum well sustaining a high mobility two dimensional electron gas (2DEG) located 37nm below the top surface and surrounded  
20 by two AlGaAs spacer layers 20. Below the structural layer stack 10 is a 400nm  $\text{Al}_{0.8}\text{Ga}_{0.2}\text{As}$  sacrificial layer 22. Sacrificial layer 22 in turn is disposed on an even thicker n+ substrate which provides a back electrode and mechanical support for chip  
28.

Fig. 1a is an energy level diagram for the heterostructure of Fig. 1b. The  
25 thickness or position, t, within stack 10 is shown on the vertical scale with the energy

level,  $\epsilon$ , in MeV on the horizontal scale. The Fermi energy  $\epsilon_F$ , is taken as the zero energy level. With the exception of a small amount of conduction in some sidebands, most of the electron conduction is confined to the 2DEG layer 18.

Note that the stack structure 10 was intentionally made asymmetric to avoid neutralizing the piezoelectric effect of GaAs layer 18, i.e. layer 18 is not in the center of the stack 10, but is fabricated to lie to one side of stack 10. As a result, layer 18 will be subjected to only tension or only compression along with the stretched or compressed layers on its side of the stack 10 as the stack is strained. The stack 10 and sacrificial layer 22 comprise the chip 28. In fact the fabrication of overlying passivating or other layers on layer 18 gives rise to a built-in strain without the imposition of external forces.

After ohmic contacts 24 are deposited, a thick layer 26 of PMMA is spun on the chip 28, followed by a single electron-beam lithography step to expose trenches 34 in PMMA layer 26 that isolate the beam 30 from its side gates 32 as shown in Fig. 2. PMMA layer 26 is then employed as a direct mask against a low voltage electron cyclotron reactor (ECR) etch performed to further etch the trenches 34 to the sacrificial layer 22. After stripping off the PMMA layer 26, the final structure of Fig. 2 relief is achieved by removing the sacrificial layer 22 beneath the beam 30 with diluted HF.

To minimize the damage to the 2DEG layer 18 from dry etching, significant efforts have been expended to optimize the etching process. After experimenting with numerous plasma mixtures, a  $\text{Cl}_2/\text{He}$  plasma was chosen because of its excellent etching characteristics such as smooth surface morphology and vertical sidewall without attacking the PMMA thus leaving a well defined mask edge. A stable etching speed at  $35\text{\AA}/\text{s}$  is obtained in an otherwise conventional ECR chamber diagrammatically depicted in cross-sectional view in Fig. 3c.  $\text{Cl}_2$  and He gas supplied at volume flow rate (sccm) ratio 1:9 respectively through orifices 202 to a plasma

chamber 200 which has been partially evacuated to 3 mTorr and the gases are ionized by 300W microwave power to etch the trenches 34 in Fig. 2 to define beam 30 while the chip 28 has 20W of constant RF power at 150 V applied to it.

The process is further illustrated in Fig. 3d. At step i the stack 10 including the  
5 quantum well structure comprised of the  $\text{Al}_{0.8}\text{Ga}_{0.2}\text{As}/\text{GaAs}$  sandwich of Fig. 1b is supplied on sacrificial layer 22. At step ii a 800 nm thick PMMA mask 26 is spun onto the surface of stack 10 and patterned using electron beam lithography to form the outline of what will become the doubly clamped beam 30 and side gates 32 (formation of the gates 32 is omitted from Fig. 3d for the sake of simplicity). At step iii the low  
10 damage ECR etch described above is performed to transfer the PMMA pattern into the underlying stack 10. At step iv a selective wet etch is performed to preferentially remove the exposed portions of sacrificial layer 22. At step v PMMA mask 26 will be stripped off using acetone or a plasma etch.

To demonstrate that the etching process does not affect the 2DEG layer 18, we  
15 have also fabricated suspended Hall effect bars with the same method and extensively characterized the suspended 2DEG that results. Before processing, the initial mobility and density after illumination are  $5.1 \times 10^5 \text{ cm}^2/\text{Vs}$ ,  $1.26 \times 10^{12} \text{ cm}^{-2}$  respectively. With our improved low damage etching, the mobility can be maintained at  $2.0 \times 10^5 \text{ cm}^2/\text{Vs}$ , while the electron density is somewhat reduced to  $4.5 \times 10^{11} \text{ cm}^{-2}$ . We observed well-  
20 developed quantum Hall plateaus in the etched structure even with channel width as small as  $0.35 \mu\text{m}$ . In longitudinal resistance measurements, we detected a low field maximum, corresponding to maximal boundary scattering when the electron cyclotron motion diameter matches the electrical width of the suspended wire. From the position of this peak, we are able to deduce the depletion to be  $0.1 \mu\text{m}$  on each side of the wire.  
25 We also confirmed ballistic behavior of electrons from transport measurement on the

Hall cross-junction. Both "last Hall plateau" and "negative bend resistance" are present in all of the devices 12. The transport mean free path was found to be approximately  $2\mu\text{m}$ .

In nanoelectromechanical (NEMS) system, both the induction and the detection of motion pose material challenges. In devices 12 of Fig. 2, the actuation is relatively easy and very effective. An RF-drive is supplied directly to one or both of the side gates 32, which is a large area of 2DEG connected to the output of a network analyzer (not shown) through an alloyed ohmic contact 24 in Fig. 1. Inducing the out-of-plane vibration of beam 30 through one or more side gates 32 is unique. Since the gate-to-beam separation,  $d$ , can be as narrow as 100 nanometers, a small driving amplitude proves sufficient. In the illustrated embodiment, all the trenches 34 have a constant width of  $0.5\mu\text{m}$ . The devices 12 are first measured at 4.2 K in vacuum. A constant DC sensing current ranging from 0 to  $26\mu\text{A}$  is supplied to the vibrating beam 30 through a 10 mH RF-choke 36, whose value is chosen big enough to avoid loss of the small signal that is induced. The oscillatory signal is picked up by a low temperature amplifier 38 placed close in proximity to the device 12, whose output is led out of the cryostat in which device 12 is immersed through a coaxial cable 39. Before connecting the signal to the input of network analyzer, a room temperature amplifier (not shown) may be used to improve the signal-to-noise ratio. The combined amplifiers have a voltage gain of about 200 in the frequency range of the illustrated experiments.

A typical completed device 12 is shown in the microphotograph of Fig. 3a and is schematically depicted in Fig. 3b. A constant DC bias current ( $I_b$ ) from current source 35 is sent through a large RF-choke 36 (about 10mH) before reaching the beam 30. Gate drive voltage applied to gate 32 consists of both DC and RF components:  $V_g = V_g^{(0)} + v_g e^{i\omega t}$ . The induced signal can be expressed as  $V = V^{(0)} + v e^{i(\omega t + \varphi)}$ , where the DC



voltage potential  $V^{(0)} = I_b R_{dc}$  is blocked by a capacitor 37, C, and the oscillating component is amplified at both liquid helium and room temperature. The beam 30 is 0.5 $\mu$ m wide and 6 $\mu$ m long, having a calculated spring constant of 0.25 N/m. When cooled to liquid helium temperature, their two-terminal resistance is about 100 k $\Omega$ . After  
 5 illumination, this drops to about 5 k $\Omega$ . The electrical width of the beam 30 is about 0.3  $\mu$ m with  $R = 170\Omega$ .

We observed very strong vibration signal around a first mechanical resonance. The magnitude response curves at various driving amplitudes are shown in Fig. 4a, which is a graph of output voltage magnitude verses frequency. Calculations confirm  
 10 that this resonance corresponds to the first out-of-plane vibrational mode, i.e. out of the plane in which the beam normal lies. When the drive amplitude is increased above 45mV, the response curve becomes nonlinear and assumes an asymmetric Lorentzian shape. In the linear response region, the amplitude at resonance is proportional to the AC gate voltage amplitude as shown in the inset graph of Fig. 4a.

15 To clarify the origin of the observed signal, we fixed the gate drive at 10 mV and then varied the DC bias current from -26  $\mu$ A to 0 then to 26  $\mu$ A. The response amplitude verses the drive amplitude at resonance is presented in Fig. 4b. Two features are evident from this data. First, at the highest currents close to 20  $\mu$ A, the signal becomes saturated for two reasons: (a) Joule-heating of the small beam 30, and (b), saturation of  
 20 the drift velocity at such high applied electrical field (about 15kV/m). Second, at intermediate current, the signal strength at resonance is proportional to the DC bias current, as indicated in the inset of Fig. 4b. In addition, when we reverse the current direction, we also find that the induced signal changes its sign (180 degree phase change).

25 Therefore we conclude that the dominant contribution to the observed signal is a

change of resistance due to beam vibration. This appears to originate from both the piezoresistive effect of bulk GaAs and transverse piezoelectric charge gating of 2DEG. Note that a small signal is observed even for zero current bias. From the slope of the linear part in the inset of Fig. 4b, a nominal drive of 10mV induces a resistance change  
 5 of about 100 in the device 12. The piezoelectric property of the beam 30 is used to induce oscillation of the beam, while its piezoresistive property is used for sensing oscillation.

We now estimate the sensitivity of this technique. By looking at the critical amplitude at the onset of nonlinearity, we can determine the amplitude of vibration of  
 10 the resonating beam 30. This critical displacement amplitude depends only on the geometry of the beam 30, and is approximately given as

$$x_c \sim \frac{2h}{\sqrt{0.5Q(1-\nu^2)}} \quad (1.1)$$

where  $h$  is the thickness of the beam in the vibration direction, and  $\nu$  is the Poisson's ratio for GaAs. Plugging in measured values of  $Q = 2600$  and  $\nu = 0.31$ , we obtain  $x_c = 6$   
 15 nm, which is attained at a drive level of about 45 mV. The minimum resolvable signal is achieved at 0.1 mV drive and about 5  $\mu$ A sensing current. Hence, at the highest possible current of 20  $\mu$ A, we can detect a resonance at  $x_c/450/4 = 0.03 \text{ \AA}$ , or  $3 \times 10^{-3} \text{ \AA}$   $/\sqrt{\text{Hz}}$ , which is consistent with our estimate based on Johnson noise from beam resistance at 4.2K. The corresponding force sensitivity is 75fN/ $\sqrt{\text{Hz}}$ , which is  
 20 comparable with previous schemes to detect small NEMS resonators or transducers by optical interferometry and the magnetomotive method. The required force to drive the beam to nonlinearity threshold is 1.5 nN. The displacement resolution can be improved by using 2DEG heterostructures with even higher mobility, or by operating at about 100 mK with a state-of-the-art low temperature preamplifier.

Note that in Figs. 4a and 4b all the driving force we applied corresponds to an applied AC gate voltage. We did not find any significant change of resonant frequency or magnitude with DC bias on the gate. This is indicative of a coupling mechanism different from electrostatic force between the gates 32 and the beam 30. Electrostatic force is proportional to the product of DC and AC components of gate potential so that the response should directly scale with the DC gate voltage. This assumes a direct Coulomb interaction between coupling plates. In our in-plane-gate configuration, the net charge on the beam is  $C (V_g^{(0)} + v_g e^{i\omega t})$  where  $V_g^{(0)}$  is the DC signal magnitude,  $v_g$  is AC signal magnitude and  $C$  is the capacitance between coplanar 2DEG areas at the gates 32 which has an estimated value of  $18 \text{ aF}/\mu\text{m}$ , which is very small compared to parallel plates. With a nominal  $1 \text{ V}$  DC gate voltage, there are only a few hundred induced electron charges on the beam 30. The upper bound of the electric field applied on the gate is  $(V_g^{(0)} + v_g e^{i\omega t})/d$ , where  $d$  is the beam-to-gate separation distance as shown in Fig. 2. Thus the total electrostatic force applied on the beam 30 with angular frequency  $\omega$  is  $f = C V_g^{(0)} v_g e^{i\omega t} y_0/d^2$  where  $y_0$  is a static offset. Only a projection of this force drives the beam along the out-of-plane ( $z$ ) direction perpendicular to the plane of the drawing of Fig. 3b. A reasonable estimate of the effective  $z$ -component of this force is,

$$f_y = C V_g^{(0)} v_g e^{i\omega t} y_0/d^2 \quad (1.2)$$

where  $y_0$  is a static offset due to, e.g., uncontrolled asymmetry of suspended beam 30. A  $10\text{nm}$  misalignment of the beam 30 with respect to gate 32 should be observable in devices 12 but was not seen. Therefore, we take this number as the upper limit of in the estimation of  $y_0$ . At a nominal  $1\text{V}$  DC gate voltage,  $45\text{mV}$  AC gate voltage, the force originating from the electrostatic drive mechanism is calculated to be  $f_y = 0.2 \text{ pN}$ . This

is four orders of magnitude smaller than the force required to drive the beam 30 into non-linear response.

For a suspended beam with strictly symmetric structural heterostructure, the static net stress is zero. Therefore in this case, the dipole-dipole actuation is a second order effect. Build-in strain in this heterostructure is induced by an intentionally designed asymmetric quantum well structure layer. Alternatively, by making bimorph structure that contains piezoelectric layer, due to the lattice mismatch of the bilayer structure, a build-in stress can develop on the beam and induce a static dipole in the beam. ( $p_2$  in Fig. 2). The piezoelectric layer could be GaAs or other III-V semiconductors, PZT, ZnO etc. The other component,  $p_1$  in Fig. 2, forms between the 2DEG layer of the side gate and a conducting substrate or the chip carrier. Given the absence of electrostatic A.C. forces, we propose that a new driving mechanism, a short-range dipole-dipole interaction, is dominant in our nanoelectromechanical system. This dipole-dipole interaction potential can be expressed as,

$$U = \int \frac{1}{4\pi\epsilon_0} \frac{p_2 dp_1}{r^3} \quad (1.3)$$

which can be understood as RF-coupling between two dipole moments  $dp_1$  and  $p_2$  as diagrammatically depicted in Fig. 2 showing a dipole charge separation 41,  $p_1$ , on beam 30 and a differential dipole charge separation 43,  $dp_2$ , on gate 32 in a differential slice  $dr$  taken perpendicular through the plane of Fig. 3b and Fig. 2. Here  $dp_1$  is the dipole momentum of a slice of the gate,  $dp_1 = \epsilon_r \epsilon_0 L v_g e^{i\omega t} dr$ , and  $p_2$  is the fixed dipole moment due to piezoelectric effect of strained GaAs/AlGaAs beam 30.  $z$  is the out-of-plane beam displacement,  $p_2 = 3E d_A w t^2 z/L$ , and  $L$ ,  $w$  and  $t$  are beam length, width and thickness as shown in Fig. 2.  $\epsilon_r$  is dielectric constant of GaAs. Here  $E$  is about

85Gpa is Young's Modulus and  $d_A$  at about 3.8 pC/N is the appropriate piezoelectric constant of AlGaAs. The resulting force along z direction is,

$$f_z = \frac{\partial U}{\partial z} = \frac{3\epsilon_r}{4\pi} (Ed_A) \left( \frac{wt^2}{d^2} \right) v_g e^{i\omega t}$$

This force is *independent* of the DC gate voltage, consistent with our observation. At 45mV AC gate voltage drive,  $f_z$  is estimated to be 1.2nN from this mechanism, four orders of magnitude higher than the direct Coulomb interaction. This is consistent with to the force we observe at the onset of non-linearity. Because of its short-range characteristics, this dipole-dipole interaction is unique to NEMS and is insignificant in microelectromechanical systems (MEMS).

We have also studied the temperature dependence of our strain sensitive devices. Measurements were performed at three different temperatures in vacuum. The results are shown in the graph of Fig. 5. The drive and sensing current are kept at the same level. The devices 12 perform exceptionally well at liquid helium and nitrogen temperatures, but at room temperature, the response is diminished. The decay of signal strength at resonance with respect to temperature can be explained by the significant reduction of 2DEG mobility at higher temperature. At elevated temperature the increased two-terminal beam resistance acts as a large voltage divider, and only a small fraction of induced signal voltage drops across the input of RF-amplifier 38.

### *Parametric Amplifier*

A nanometer-scale mechanical parametric amplifier is provided based purely upon the intrinsic mechanical nonlinearity of a doubly-clamped beam. Operating in degenerate mode, a parametric modulation of the beam's force constant at twice the signal frequency is produced by the application of an alternating longitudinal force to its

ends. This provides stable, nearly thousand-fold small-signal mechanical gain at the threshold for parametric oscillation. For large signals, we find the gain saturates below this threshold; in this regime the device performs as a limiting preamplifier. At the highest gains noise-matched performance at the thermodynamic limit is achieved. A simple theoretical model explains the observed phenomena and indicates that this approach offers great promise for achieving output-coupled quantum-limited nanoelectromechanical systems.

The parametric amplifier described in the illustrated embodiment as shown in the microphotograph of Fig. 6 operates on a suspended nanomechanical transducer or beam 30 with a natural frequency at 17 MHz, with a gain-bandwidth product of 2.6 kHz, and requires pump voltages of only a few mV and power on the order of 1  $\mu$ W to yield small-signal gain approaching 1000. The modulation of the spring constant is purely mechanical, requiring no capacitor plate as in the prior art, and precisely controlled by the fabrication geometry, requiring no prestress as in the prior art. The mechanism employed in the illustrated embodiment permits high gain-dynamic range product, in excess of 65 dB. Phase dependent amplification of thermomechanical fluctuations is observed at 4 K. Due to the stiffness of the device 40, detection sensitivity is limited by noise in the electrical readout amplifier 38, and is insufficient to observe thermomechanical noise. However, the device 40 is operated as a mechanical preamplifier, demonstrating a dramatic improvement in signal-to-noise ratio for small-amplitude harmonic motion.

Device 40 in Fig. 6 was fabricated by electron beam lithography from an epitaxial layer of silicon carbide on a silicon substrate. The device 40 was patterned by a vertical plasma etch of the silicon carbide layer and suspended by an isotropic plasma etch removing the supporting silicon. Device 40 is comprised of a signal beam 31 supported

at either end by perpendicular pump beams 42 as shown in the microphotograph of Fig. 6. The lateral extent of the device 40 is 17.5  $\mu\text{m}$  and its thickness is 200 nm. The device 40 is measured in vacuum at 4 K in a magnetic field of  $B = 8$  Tesla perpendicular to the chip's surface, so the Lorentz force provides an excitation of the signal beam 31, and the magnetomotive technique is used to detect its motion.

The spring constant of the signal beam 31 is modulated by the application of an alternating current  $I$  flowing through path 44 at a frequency  $2\omega_0$  through the pump beams 42 as shown in the diagram of Fig. 7, where  $\omega_0$  is the fundamental frequency of beam 31,. The Lorentz force,  $T$ , generated by this current applies sinusoidal compression and tension to the signal beam 31 is:

$$T = 2 B I L_2 \zeta \cos(2\omega_0 t) \quad (2.1)$$

where  $L_2$  is the length of the pump beams 42 and  $\zeta$  is a geometric factor to account for the finite restoring force of the pump beams 42. In principle,  $\zeta$  can be evaluated from a finite element simulation. The longitudinal force perturbs the flexural spring constant for in-plane motion of the signal beam 31 with an amplitude of:

$$\frac{k_p}{k_1} = \frac{12}{\pi^2 E t_1 w_1} \left(\frac{L_1}{w_1}\right)^2 T \quad (2.2)$$

where  $E$  is Young's modulus, and  $w_1, L_1$ , and  $t_1$  are the width, length and thickness of the signal beam 31.

For small displacements, the equation of motion of the signal beam 31 under the influence of the pump and a harmonic excitation  $F_a$  is:

$$m\ddot{x} + \frac{m\omega_0}{Q} \dot{x} + (k_1 + k_p \cos(2\omega_0 t)) x = F_0 \sin(\omega_0 t + \phi) + f_n \quad (2.3)$$

where  $m$  is the effective mass,  $Q$  is the quality factor, and  $f_n$  is a thermomechanical noise. Above a threshold pump amplitude

$$k_1 = 2 \frac{k_1}{Q} \quad (2.4)$$

5

the gain of the parametric amplifier diverges. For pump amplitudes below threshold, the mechanical gain depends on the relative phase  $\Phi$  between the excitation and the pump:

$$G(\phi) = \left[ \frac{\cos^2 \phi}{(1 + \frac{k_p}{k_i})^2} + \frac{\sin^2 \phi}{(1 - \frac{k_p}{k_i})^2} \right]^{\frac{1}{2}} \quad (2.5)$$

10

Although this expression for gain diverges as  $k_p$  approaches the threshold, in practice nonlinearities in the system cause the gain to saturate. The dominant nonlinearity in our system is the geometric stiffening due to flexure, which results from longitudinal stretching of the signal beam 31 clamped by semi-rigid supports. To develop a model for saturation, we incorporate the cubic stretching term into the equation of motion:

15

$$m\ddot{x} + \frac{m\omega_0}{Q}\dot{x} + (k_1 + k_{p0} \cos(2\omega_0 t))x + k_3 x^3 = F_0 \sin(\omega_0 t + \phi) + f_n \quad (2.6)$$

20

where



$$k_3 = 0.36k_1 / t^2 \quad (2.7)$$

If we consider motion at the fundamental frequency, with phase  $\Phi = 0$  chosen for  
 5 maximum gain  $G$ , then

$$x = G x_0 \sin(\omega_0 t) \quad (2.8)$$

and the cubic term perturbs the spring constant at  $2\omega_0$  to oppose the action of the  
 10 pump:

$$x^3 = -\frac{G^2 x_0^2}{2} \cos(2\omega_0 t) x + \frac{G^2 x_0^2}{2} x \quad (2.9)$$

Ignoring the linear term above, we derive an equation for the steady-state  
 amplitude of motion  $x = Gx_0$  :

15

$$\frac{k_3^2}{4} x^5 + k_3 (k_t - k_{p^0}) x^3 + (k_t - k_{p^0})^2 x - k_t^2 x_0 = 0 \quad (2.10)$$

The response of the parametric amplifier is measured with the circuit shown in  
 the schematic diagram of Fig. 8. The lengths of the coaxial cables 46 and 48 to the  
 20 pump beams 42 and signal beam 31 are chosen so that they act as 1-1 impedance  
 transformers at  $2\omega_0$  and  $\omega_0$  respectively. Pump beams 42 are coupled through cable  
 46 to a driving oscillator 50 operating at  $2\omega$  and an equivalent thermoelectric noise  
 source 60. A virtual output oscillator 52 operating at  $\omega$  is coupled through load  
 resistance 54 through cable 48 to signal beam 31 and comprises an output reference

signal indicative of the parametric oscillation of signal beam 31. The output from signal beam 31 is coupled through amplifier 56 to a display or measurement device 58. The electrical response is then the superposition of the mechanical motion on the baseline electrical resistance of the signal beam 31. To determine the mechanical gain we compare the electrical response on and off resonance, as measured by a spectrum analyzer:

$$G = \frac{V_{\text{on resonance}}^{\text{pump on}} - V_{\text{off resonance}}^{\text{pump on}}}{V_{\text{on resonance}}^{\text{pump off}} - V_{\text{off resonance}}^{\text{pump off}}} \quad (2.11)$$

In order to verify the effectiveness of the pump, we substituted a network analyzer for the signal source 50 and spectrum analyzer 58 in the schematic in Fig. 8, and measured the frequency shift of the resonance peak as a function of DC pump force. From the fit in Fig. 9, we find the frequency shift to be  $\Delta f/f = 1.59 / \text{mN}$ , neglecting the finite restoring force of the pump beam 42 (assuming  $\zeta = 1$  in (1)). From equation (2.2), the expected variation is  $\Delta f/f = 6.24 / \text{mN}$ . The discrepancy between these values suggests that the stiffness of the pump beam 42 does indeed reduce the effective pump force applied to the signal beam 31. To evaluate  $\zeta$ , we conducted a finite element mechanical simulation of the structure in which static forces totaling 1 nN were applied transverse to the pump beams 42 as shown in the diagram of Fig. 10. From the calculated compression, 87 pm, of the signal beam 31 in the model, the effective compressive force applied to the signal beam 31 can be found:

$$T = E t w \Delta x/x \quad (2.12)$$

Thus, we find  $T = 0.235$  nN and  $\zeta = 0.235$ , so our measured frequency shift is actually  $6.77$  /mN. The agreement with the expected value indicates that our model  
 5 accounts for the effectiveness of the pump.

To further demonstrate that the observed parametric effect is due to the Lorentz force on the pump beam 42, Fig. 11 is a graph which shows a measurement of the phase-dependent gain of the amplifier of Fig. 8 at two different magnetic fields. The signal beam 31 is driven at its fundamental frequency  $\omega_0$  and the pump beams 42 are  
 10 driven at  $2\omega_0$ , referenced to the signal beam 31 through a variable phase shifter (not shown). The motion of the signal beam 31 is either amplified or deamplified, depending on the phase difference between the motion of the signal beam 31 and the excitation of the pump beams 42. As equation (2.5) predicts, the maximum gain occurs at  $\Phi = \pi/2$  and the minimum gain occurs at  $\Phi = 0$ . As equations (2.1) and (2.2) predict, for  
 15 stronger magnetic field, the pump-induced frequency shift is greater, so the maximum gain is greater and the minimum gain is smaller. Although amplification and deamplification are greatest when the pump is at exactly  $2\omega_0$ , substantial variation in gain is possible off-resonance. For excitation at  $\omega$  slightly off-resonance, two sidebands are created, one at  $\omega$  and one at  $2\omega_0 - \omega$ . Fig. 12 shows the dominant sideband  $\omega$  of  
 20 the signal beam's response to a fixed excitation with the phase shift set for maximum gain. At high gain, the action of the pump substantially reduces the bandwidth of the resonance. For a pump voltage of  $8.2$  mV, the bandwidth is reduced from  $1760$  Hz to  $35$  Hz.

As the pump amplitude approaches the threshold value, the gain of the  
 25 parametric amplifier on resonance is expected to increase dramatically. When our

device is operated just below threshold, at 8.2 mV, the gain of 39 at  $\Phi = 0$  is sufficient to observe amplified thermomechanical fluctuations, as shown in the graph of Fig. 13. The response of the signal beam 31 to thermomechanical fluctuations has a Lorentzian line shape which is narrowed by the parametric amplifier. Since the fluctuating force is not coherent with the pump, the gain for this peak should be averaged over phase. Assuming an average gain of 39, the amplitude of the peak corresponds to an rms amplitude of motion of  $550 \text{ fm/Hz}^{1/2}$ , or  $14 \text{ fm/Hz}^{1/2}$ . The amplitude of thermomechanical fluctuations for a simple harmonic oscillator on resonance is given by

$$\sqrt{S'_x} = \sqrt{\frac{4k_B TQ}{\omega_0 k}} \quad (2.13)$$

where the spring constant  $k = m\omega_0^2$ ; is 32 N/m, which yields a value of  $26 \text{ fm/Hz}^{1/2}$  for the signal beam 31. The discrepancy between the values can be attributed to an error due to the average gain approximation and in the calculation of spring constant.

We observed the phase dependence of the amplified thermomechanical fluctuations by replacing the spectrum analyzer with a radiofrequency lock-in amplifier (not shown) referenced to the pump beams 42. As Fig. 14 demonstrates, for pump voltage close to threshold, the fluctuations are clearly amplified, but in one quadrature only (i.e. the phase relationship between  $\omega$  and  $2\omega$ ). No effect is observed in the other quadrature, because the total noise in that quadrature is dominated by the phase-independent noise at the input of the linear electrical amplifier 56.

Just as the Brownian motion of the signal beam 31 is amplified without the addition of mechanical noise, so too is harmonic motion. Since in our system, the electrical amplifier dominates the noise level, the signal-to-noise ratio for measurement

of harmonic motion of the signal beam 31 can be improved dramatically by parametric amplification. The graph of Fig. 15 compares the gain for a harmonic excitation which would yield an rms amplitude of motion of 1.2 pm with the overall noise level in each quadrature. Near the threshold pump amplitude, the signal-to-noise ratio improves by a factor of nearly 100 with respect to the  $\Phi = \pi/2$  quadrature. As a consequence of the dominance of the electrical amplifier's input noise over thermomechanical noise, the signal-to-noise ratio is also improved in the  $\Phi = 0$  quadrature, although to a lesser extent. This result describes the most fundamental application of a parametric amplifier, namely as a mechanical preamplifier.

The dynamic range of the amplifier is critically important to this application. For a harmonic excitation of 47 fm in the absence of a pump, our device exhibits gain as high as 800, as shown in the graph of Fig. 16. However, for larger excitations, the gain saturates at much lower values. Fig. 16 clearly demonstrates that the point at which the gain begins to saturate depends solely on the amplitude of motion, not the excitation. The saturation begins at an rms amplitude of -360 pm, and gives a good approximation to the upper bound for the dynamic range of the amplifier. Ultimately, the upper limit of dynamic range is a direct consequence of nonlinearities in the system. In our system, the dominant nonlinearity is expected to be the cubic term in the expansion of the flexural spring constant.

### *Sensitivity of Piezoresistive NEMS Displacement Transducers in Vacuum*

One of the most important engineering challenges to be faced is optimization of the readout system that measures a NEMS cantilever displacement. Shown in the SEM photograph of Fig. 51 is an example of a practical device, a cantilever 190, which incorporates a piezoresistive strain transducer. The transducer converts the motion of

the cantilever 190 into an electrical signal, in this case via the strain-induced change in resistance of a conducting path patterned from p+ doped Si epilayer disposed on the top surface of the cantilever 190. For the purposes of illustration the bioNEMS transducer or cantilever 190 shown in perspective view in the microphotograph Fig. 51 can be analogized as having the form of "a diving board with a cutout at its base". The geometry of the device 190 causes dissipation to occur predominantly within a constriction region 192 comprised of one or more legs 194 of width  $b$ , which region 192 allows for enhanced or variably designed flexural stiffness of cantilever 190. It is also to be understood that cantilever 190 will have conventional electrodes (not shown) provided whereby a conventional external measurement circuit (not shown) providing a bias current may measure the change in piezoresistivity of legs 194 as they flex. In addition, an external driving force may or may not be applied in a conventional manner to cantilever 16 depending on the application and design choice. In the preferred embodiment there are two legs 194. We assume that a temperature rise of order 10 K is tolerable at the biofunctionalized tip 196 of the cantilever 190, which has a length,  $l$ , a width,  $w$ , and a thickness,  $t$ , resonant frequency in vacuum  $\omega_0/2\pi$  and force constant  $K$ .

We characterize the transducer's performance by its responsivity,  $R$ , in units volts/m,  $R = I G$ , where  $I$  is the bias current, while  $G = \frac{\partial R_T}{\partial x}$  and  $R_T$  are the gauge factor and the two-terminal resistance of the transducer, respectively.

Near resonance, the force spectral density of thermomechanical displacement fluctuations is given by  $S_V^r = 4k_B T \gamma = 4Kk_B T / Q$ . Three additional terms arising from electrical noise in the readout process must also be included. These must be referred back to the input, *i.e.* to the displacement domain using the factor  $1/R^2$ . The first arises from the thermal voltage noise of the piezoresistive transducer,  $S_{VT} = 4k_B T R_T$ , while the

second arises from the readout amplifier's voltage and current noise,  $S_{VA} = S_V + S_I R_T^2$ ; where  $S_V$  and  $S_I$  are the spectral density of the amplifier's voltage and current noise, respectively.

The sum of these fluctuations yields what we term the total coupled  
5 displacement noise, which is the actual displacement sensitivity of the entire system,

$$S_x^{(C)} = S_x^{(\gamma)} + \frac{1}{R^2} \{S_{VT} + S_{VA}\}$$

From this we can determine the coupled force sensitivity of the electromechanical system, which at resonance is given by

$$10 \quad S_F^{(C)} = K^2 S_x^{(C)} / Q^2$$

where  $K$  is the spring constant and  $Q$  is the quality factor of the cantilever beam.

The somewhat complex mechanical device shown in Fig. 51, compared to a simple cantilever geometry, provides a higher degree of compliance while providing low mass (if its overall size is kept small). Its spring constant is more complex than that of  
15 the simple cantilever, and can be written as \*\*\*EQN corrected\*\*\*

$$K = \frac{Et^3}{\frac{4l^3}{w} + (2l_1^3 - 6l_1^2 l_2 + 6l_1^2 l_3) \left( \frac{1}{b} - \frac{2}{w} \right)} \quad 7.1$$

The variables characterizing the device geometry are depicted in Fig. 51 and are summarized for the cantilevers discussed here in Table 3 which shows the physical  
20 parameters for three prototype Si nanocantilevers. The parameters tabulated are thickness,  $t$ ; width,  $w$ ; length,  $l$ ; constriction width,  $b$ , and length  $l_1$ ; frequency in vacuum  $\omega_0 / 2\pi$ ; force constant  $K$ ; and resistance  $R_T$

Table 3

cantilever	$t$	$w$	$\ell$	$\ell_1$	$b$	$\omega_0 / 2\pi$	$K$	$R_T$
1	130nm	2.5 $\mu$ m	15 $\mu$ m	4.0 $\mu$ m	0.6 $\mu$ m	0.51MHz	34mN/m	15.6k $\Omega$
2	110nm	900nm	6 $\mu$ m	3 $\mu$ m	300nm	3.1MHz	145mN/m	22k $\Omega$
3	37nm	300nm	2 $\mu$ m	1 $\mu$ m	100nm	9.2MHz	48mN/m	67k $\Omega$
4	30nm	30nm	0.3 $\mu$ m	20nm	10nm	360MHz	1.0N/m	16k $\Omega$

In the illustrated device, cantilever 190 has thickness 130 nm, with the topmost 30nm comprised of a heavily (p+) doped Si epilayer, while the remaining 100nm is intrinsic Si layer underlying the Si epilayer. The piezoresistive transducer is patterned from the p+ boron doped Si ( $4 \times 10^{19}/\text{cm}^3$ ) with the current path in legs 194 oriented along the  $\langle 110 \rangle$  direction. The gauge factor for this cantilever is given by

$$G = \frac{3\beta\pi_L K(2l - l_1)}{2bt^2} R_T \quad 7.2$$

The parameter  $\pi_L$  is the piezoresistive coefficient of the p+ transducer material ( $4 \times 10^{-10} \text{ m}^2/\text{N}$  for p-type  $\langle 110 \rangle$  silicon). The parameter  $\beta$ , a coefficient between 0 and 1, is used to account for the finite thickness of the conducting layer.  $\beta$  approaches 1 monotonically as the carriers become confined to a surface of infinitesimal thickness. We expect  $\beta \sim 0.7$  for our cantilevers. The factor  $\beta$  accounts for the decrease in  $G$  due to the finite thickness of the conducting layer;  $\beta$  approaches unity as the carriers become confined to a surface layer of infinitesimal thickness. We shall assume  $\beta = 0.7$  for our epilayer. For the cantilever pictured in Fig. 51 we observe,  $G = 3.3 \times 10^7 \Omega/\text{m}$ . For the transducer geometry pictured, a two-terminal (equilibrium) resistance of  $R_T = 15.6 \text{ k}\Omega$  is obtained. Note that this implies  $G/R_T \sim 2.1 \text{ ppm}/\text{nm}$ .

We now investigate what are the constraints upon the level of current bias



applied in the circuit in legs 194. The force sensitivity attainable clearly hinges on the maximum level of bias current that is tolerable, given that the responsivity is proportional to bias current,  $\mathcal{R} = I/G$ . The largest practical level is determined by the maximum temperature rise deemed acceptable. The geometry of the prototype devices

5 causes dissipation to occur predominantly within the constriction regions 192 of width  $b$ . We assume that a maximal temperature rise of order 10K is tolerable. We treat the problem as one dimensional, with the constricted region 192 of beam 190, of length  $l_1$  and cross sectional area  $A$ , heat sunk at the supporting end 195. It is assumed that all

no heat is exchanged through the vacuum. In the dissipative region  $x < l_1$  we have

10  $2\kappa_{Si}tl_1b \frac{d^2T}{dx^2} = -I^2R$ , where  $\kappa_{Si}=1.48 \times 10^2$  W/mK at 300K is the thermal conductivity of silicon. Assuming the dissipation in the region of cantilever 190 beyond the constriction region 192 to be negligible, we apply the boundary condition  $dT/dx = 0$  at  $x = l_1$ . This simple thermal conductance calculation indicates that a maximum temperature rise of 10K is obtained with a steady-state bias current  $I = 60\mu A$ , leading to a power dissipation

15 of roughly 60 $\mu$ W. For this bias current, our prototype device yields a responsivity  $\mathcal{R} = I/G \sim 2\mu V/nm$ .

With knowledge of these parameters we can now estimate the coupled force sensitivity of the prototype system. For cantilever 190, starting at room temperature and assuming that a 10K temperature rise, the transducer-induced thermal voltage noise,

20 referred to the force domain is found to be  $K\sqrt{S_{VT}}/( \mathcal{R} Q) = 92 \text{ aN}/\sqrt{\text{Hz}}$  at resonance for a  $Q$  of 2000. For a typical low noise readout amplifier with voltage and current noise levels (referred to input) of  $\sim 4\text{nV}/\sqrt{\text{Hz}}$  and,  $\sim 5\text{fA}/\sqrt{\text{Hz}}$ , respectively, these same parameters yield an amplifier term  $K\sqrt{S_{VA}}/Q = 23\text{aN}/\sqrt{\text{Hz}}$ . For this cantilever, the force spectral density of the thermomechanical displacement fluctuations is  $S_F^y = 300 \text{ aN} /$

$\sqrt{\text{Hz}}$ . The total transducer noise for improved quality factors are given in Table 4.

Clearly the noise from thermomechanical displacement fluctuations is dominant. This may be decreased by reducing the dimensions to increase the resonance frequency and decrease the spring constant.

5 To illustrate the benefits of further scaling downward in size for such devices, we consider two smaller cantilevers having geometry identical to that of Fig. 51, but with  $l = 6\mu\text{m}$ ,  $t = 110\text{ nm}$ ,  $w = 900\text{ nm}$ ,  $b = 300\text{ nm}$ , and  $l_1 = 3\mu\text{m}$ . Assuming this device is constructed from the same epilayer thickness ratio as cantilever 190, this yields  $R_T = 19\text{ k}\Omega$  and  $G = 2.9 \times 10^9\text{ }\Omega/\text{m}$  (cantilever #2 in Table 4).

10 For cantilever #2 we again assume that a 10K rise at the tip is tolerable. For  $Q = 2000$ , we find the transducer-induced force noise is  $K\sqrt{S_{VT}}/l(RQ) = 7.1\text{ aN } \sqrt{\text{Hz}}$ , whereas the readout amplifier's contribution, referred to the force domain, is  $K\sqrt{S_{VT}}/l(RQ) = 1.5\text{ aN } \sqrt{\text{Hz}}$ . The force spectral density of the thermomechanical displacement fluctuations is  $S_F^Y = 249\text{ aN } \sqrt{\text{Hz}}$ . For  $Q = 30000$  the force spectral density of the  
15 thermomechanical displacement fluctuations is  $S_F^Y = 64\text{ aN } \sqrt{\text{Hz}}$ .

Another device considered, "cantilever #3" is identical to cantilever #2, but is uniformly reduced in all dimensions by a factor of  $\sim 3$ . For this device  $R_T = 67\text{ k}\Omega$ , and  $G = 3.0 \times 10^{10}\text{ }\Omega/\text{m}$ . Once again using  $Q = 2000$ , this yields transducer-induced Johnson force noise  $S_F^Y = 1.5\text{ aN } \sqrt{\text{Hz}}$  and an amplifier contribution, referred to the force  
20 domain,  $K\sqrt{S_{VT}}/lQ = 0.18\text{ aN } \sqrt{\text{Hz}}$ . The force spectral density of the thermomechanical displacement fluctuations is  $S_F^Y = 83\text{ aN } \sqrt{\text{Hz}}$ . For  $Q = 30000$  the force spectral density of the thermomechanical displacement fluctuations is  $S_F^Y = 21\text{ aN } \sqrt{\text{Hz}}$ . Force sensitivity for other amounts of heating allowed is given in Table 4

Table 4

Coupled force sensitivity at resonance for room temperature piezoresistive detection

Cantilever 1		Cantilever 2		Cantilever 3		Cantilever 4	
Q	$S_F^C$ [aN/ $\sqrt{\text{Hz}}$ ]	Q	$S_F^C$ [aN/ $\sqrt{\text{Hz}}$ ]	Q	$S_F^C$ [aN/ $\sqrt{\text{Hz}}$ ]	Q	$S_F^C$ [aN/ $\sqrt{\text{Hz}}$ ]
2000	329	2000	280	2000	86	2000	62
10000	136	10000	125	10000	38	10000	28
30000	77	30000	72	30000	22	30000	16
100000	42	100000	40	100000	12	100000	8.8

For a doping density of  $4 \times 10^{19}/\text{cm}^3$  the depletion length is on the order of 2nm

5 therefore while cantilever #3 lies entirely within the domain of feasibility, it is not realistic to continue pushing the thickness any thinner than 30nm. In order to achieve the 364 MHz of cantilever #4 the length was therefore decreased without significantly decreasing the thickness. Despite the limitation imposed by the increased spring constant, the force sensitivity remains excellent with a coupled force sensitivity of  $S_F^Y =$

10 62 aN  $\sqrt{\text{Hz}}$  for a quality factor of 2000 at room temperature. The sensitivity of all four cantilevers for a variety of quality factors are summarized in table 4. Table 5 gives the analogous data for 4K.

Table 5

Coupled force sensitivity at resonance for piezoresistive detection at 4K

Cantilever 1		Cantilever 2		Cantilever 3		Cantilever 4	
Q	$S_F^C$ [aN/ $\sqrt{\text{Hz}}$ ]	Q	$S_F^C$ [aN/ $\sqrt{\text{Hz}}$ ]	Q	$S_F^C$ [aN/ $\sqrt{\text{Hz}}$ ]	Q	$S_F^C$ [aN/ $\sqrt{\text{Hz}}$ ]
2000	74	2000	22	2000	5.7	2000	16

10000	20	10000	9.8	10000	2.5	10000	4.3
30000	9.9	30000	5.7	30000	1.5	30000	2.1
100000	5.0	100000	3.1	100000	0.81	100000	1.0

### *NEMS-based Piezoresistive Force Sensing*

Force sensitivity for a piezoresistive detector is discussed above at both room  
 5 temperature and 9K. The pressure dependence of the room temperature force  
 sensitivity is also discussed. A direct measurement of the gauge factor was obtained  
 using an atomic force microscopy [AFM] to move the cantilever tip a known amount as  
 shown in Fig. 51a. This yields a direct measurement of  $G = dR_T/dx = 3 \times 10^7 \Omega/m$  and we  
 calculate  $G \sim 6 \times 10^8 \Omega/m$  for  $\beta=0.7$ . The discrepancy is attributed to diffusion during  
 10 processing. In particular, to mask these particular devices for which the membrane from  
 which the cantilever was patterned was formed using a KOH etch during early  
 processing steps, silicon nitride was grown by LPCVD at 850°C at which temperature  
 diffusion is a concern, decreasing  $\beta$  from the expected value; this high temperature  
 masking step is not necessary if a DRIE etch is used as an alternative to the KOH etch  
 15 for membrane formation.

Near resonance, the force spectral density of thermomechanical fluctuations is  
 given by

$$S_F^{\gamma} = 4k_B T \gamma = 4Kk_B T / (2\pi Q f_0) \quad 8.1$$

where  $\gamma$  is the damping coefficient in kg/s,  $f_0$  is the resonance frequency and  
 20  $Q = m f_0 / \gamma$  is the quality factor.

Near resonance, the voltage spectral density for the thermomechanical  
 fluctuations is therefore given by

$$S_V^\gamma = \frac{S_F^\gamma G^2 / l^2}{16\pi^2 m^2 f_0^2 [4(f - f_0)^2 + f_0^2 / Q]} \quad 8.2$$

Measured at the amplifier this will give

$$S_V^{\text{measured}} = \frac{S_V^\gamma}{R_T} \left( \left( \frac{1}{R_T} + \frac{1}{R_{\text{bias}}} + \frac{1}{R_{\text{amp}}} \right)^2 + 4\pi^2 f^2 C^2 \right)^{-\frac{1}{2}} + S_V^{\text{measured}} + S_V^A \quad 8.3$$

where  $R_{\text{bias}}$  is the impedance of the bias resistor, connected in parallel to the  
 5 sample,  $R_{\text{amp}}$  is the input impedance of an amplifier (not shown) to which cantilever 190  
 is electrically coupled, and  $C$  is the input capacitance of the amplifier,  $S_V^{\text{measured}}$  is the  
 Johnson noise as measured at the input to the amplifier and  $S_V^A$  is the voltage spectral  
 density from the amplifier.

Fig. 52 shows the resonance peak for the thermomechanical noise in vacuum at  
 10 room temperature, for a device of dimensions comparable to that used for the  
 measurement of the gauge factor above. The sample resistance is 16.7k $\Omega$  and is in  
 parallel with a 1 0.5k $\Omega$  resistor. The input capacitance of the amplifier is 33pF and input  
 resistance 100k $\Omega$ . We therefore expect a background from Johnson noise of 5.7nV/ $\sqrt{\text{Hz}}$   
 at 605.5kHz. The preamplifier noise was measured to be 2.5nV/ $\sqrt{\text{Hz}}$  at this frequency.  
 15 Giving a combined expected background of 6.2nV/ $\sqrt{\text{Hz}}$ . The measured background was  
 9. 13nV/ $\sqrt{\text{Hz}}$ . For this cantilever, the measured resonance frequency was 605.5kHz.  
 The measured quality factor in vacuum was 550. From Eqn. 8.1, the force spectral  
 density from thermomechanical fluctuations is therefore 1.5 fN/ $\sqrt{\text{Hz}}$ . We may invert  
 Eqns. 8.2 and 8.3 and use the Lorentzian fit to the experimental data to give a  
 20 measurement of the gauge factor, giving  $G=1.0 \times 10^8 \Omega/\text{m}$ .

The inset to Fig. 52 shows the pressure dependence of the quality factor for this  
 device. The pressure clearly has a dampening effect above 200mTorr.

Fig. 53a shows the resonance peak for the same device placed in a liquid helium cryostat. A bias current of 48 $\mu$ A was used, it is estimated that the maximal heating at this temperature (occurring at the device tip) should be given by  $I^2 R / (4 \kappa_{Si} t_b) \sim 4K$ .

So the temperature at the device tip is  $\sim 9K$ . At this temperature the resonance

- 5 frequency is 552kHz and a quality factor of  $2.1 \times 10^3$  was obtained. The force sensitivity is given by equation 8.1. Using the measured quality factor and the estimated temperature of 9K this gives a force sensitivity of 113 aN/ $\sqrt{Hz}$ . From Eq. 8.2 it is possible to extrapolate the gauge factor. This gives a gauge factor of  $1.6 \times 10^8 \Omega/m$  or an increase by a factor of 1.6 from the room temperature value, arising from an
- 10 increase in the piezoresistive coefficient increases with decreasing temperature.

Fig. 53b shows the same data for another device of the same dimensions fabricated simultaneously on the same chip. The resistance of this cantilever is 14.4k $\Omega$ . The resonance frequency of this cantilever was 620kHz and a quality factor of  $2.11 \times 10^3$  was measured. From Eq. 8.1 this gives a force sensitivity of 126 aN/ $\sqrt{Hz}$ .

15

### *Scaling of Piezoresistive Sensors*

Piezoresistors are designed to have a thin heavily doped silicon layer on top of nominally intrinsic silicon. As the devices are scaled to smaller dimensions, the effect of the depletion layer in the thin silicon layer becomes increasingly significant. The carrier

20 distribution is computed below by iterating between two procedures until convergence is attained. The first procedure adjusts the Fermi level until charge neutrality is attained. The second procedure calculates the bending of the valence band according to the equation

$$\frac{d^2 E_v}{dz^2} = \frac{e \rho(x)}{\epsilon} \quad 3.1$$

where  $E_v$  is the energy of the valence band,  $e$  is the charge of the electron,  $p$  is volume density of carriers, and  $\epsilon$  is the dielectric constant.  $p(x)$  is the charge density, given by Fermi statistics,  $\rho(x) = e(p(x) - n(x))$  where  $p(x) = 1.04 \times 10^{25} e^{-\beta(E_F - E_v)} / m^3$  is the density of positive carriers. Subject to the boundary condition:

$$5 \quad \left. \frac{d^2 E_v}{dz^2} \right|_{z=0} = \frac{e\sigma}{\epsilon} \quad 3.2$$

where  $\sigma$  is the surface carrier density. The density of surface states  $\sigma$  for equation 3.2 and 3.3 were estimated based on published values for interface state density at a silicon-silicon dioxide interface.

Setting the Fermi level to achieve charge neutrality assures that at the lower  
10 surface the boundary condition is attained,

$$\left. \frac{d^2 E_v}{dz^2} \right|_{z=t} = \frac{e\sigma}{\epsilon} \quad 3.3$$

where  $z = t$  is lower surface of the nominally intrinsic silicon, which is also normally the lower surface of the transducer or cantilever. ( $z$  is the out of plane direction)

The Fermi level,  $E_F$ , is set by the condition that charge neutrality be maintained;

$$15 \quad \int_0^t (\rho(x)/e + N_A^-(x)) dx = 0 \quad 3.6$$

where

$$N_A^-(x) = \frac{\#dopants}{\frac{1}{2} e^{-\beta(E_A - (E_F - E_v))}} \quad 3.7$$

is the density of ionized acceptor sites,  $E_A$  is the energy of the ionized acceptor sites.

$$p(x) = 1.04 \times 10^{25} e^{-\beta(E_F - E_v)} / m^3$$

$$20 \quad n(x) = 2.8 \times 10^{25} e^{-\beta(E_c - E_F)} / m^3$$

where  $\beta$  is  $1/kT$  and  $E_C$  is the energy of conduction band. Equations 3.1 and 3.6 were solved iteratively until convergence was attained.

Fig. 17 shows the carrier distribution for a sample of 130nm thickness in which the dopant layer is 30nm thick and the dopant concentration is  $4 \times 10^{25} \text{ m}^{-3}$ . The carrier  
 5 distribution for a sample of 30nm thickness for which the thickness of the doped layer is 7nm is shown in Fig. 18. In both cases the carriers are well confined.

It is clear from Fig. 18 that we are now approaching the minimum thickness that may be achieved with a conventional 2 layer structure such as cantilever 190. Further direct reduction of the dimensions is not possible without sacrificing performance as the  
 10 depletion layer thickness becomes significant relative to the dimension of the doped region. A new technique is therefore required.

#### *Confining Carriers in piezoresistive NEMS sensors*

Carrier confinement can be substantially increased by confining the carriers in a  
 15 quantum well structure as depicted in Figs. 54 a, 54b, and 54c. In these figures the conduction/piezoresistive sensing takes place in the quantum well (QW) layer 300 and the layer 302 referred to as the "confining layer" serves to confine the carriers to the QW layer 300. To accomplish this, the confining layer 302 must have a significantly lower valence band edge in the case of a p type sensor or a significantly higher  
 20 conduction band edge in the case of an n type sensor. A difference in band edge energy on the order of 0.4eV or greater is considered significant for the purposes of good carrier confinement.

In a concrete realization of the structure depicted in Fig. 54b, the top and bottom confining layers 302 and 304 might be intrinsic silicon grown in the (100) plane. While  
 25 the quantum well layer 300 could be p doped germanium.(also grown in the (100) plane



which may be epitaxially grown on the silicon layer; boron, indium and gallium are examples of p dopants in germanium). The piezoresistive sensor could then be patterned in the  $\langle 110 \rangle$  direction. The piezoresistive coefficient for  $\langle 110 \rangle$  oriented p-type germanium is 50% larger than that for silicon oriented in the same direction. The  
5 valence band edge in germanium is 0.46eV above that for silicon, which is sufficient to confine the carriers for this application.

These materials might also be used for realization of the piezoresistive layer of Fig. 54c with germanium once again serving as the quantum well 300 and intrinsic silicon as the lower layer 302.

10 This specific example of materials that may be used is not intended to limit the invention in any way. The field of confining carriers to 2DEGs or quantum wells is well developed and all of the knowledge and technology of the field may be used in the production of sensors such as those described herein.

Structures such as that described above will also have limits on the minimum  
15 thickness that may be attained both due to carrier confinement (due to finite well depth) and practical fabrication issues (due to the multiple layers). Even greater reduction in sensor thickness may be attained through the use of an insulator as the supporting layer such as shown in Fig. 55. Examples of materials that may be used for the insulating layer 306 are silicon dioxide or silicon nitride but this invention includes any  
20 insulator and should not be limited to these two.

The benefits of increased sensitivity as the thickness of a piezoresistive sensor is decreased was demonstrated in the section *Sensitivity of Piezoresistive NEMS Displacement Transducers in Vacuum*. The inventions described herein allow for a decrease in thickness beyond that which may be achieved using a conventional 2 layer  
25 structures of heavily doped silicon on intrinsic silicon.

### *Balanced Electronic Displacement Detection for VHF NEMS*

A broadband radio frequency (RF) balanced bridge technique for electronic detection of displacement in nanoelectromechanical systems (NEMS) uses a two-port  
5 actuation-detection configuration, which generates a background-nulled electromotive force (EMF) in a DC magnetic field that is proportional to the displacement of the NEMS transducer. The effectiveness of the technique is shown by detecting small impedance changes originating from NEMS electromechanical resonances that are accompanied by large static background impedances at very high frequencies (VHF). This technique  
10 allows the study of experimental systems such as doped semiconductor NEMS and provides benefits to other high frequency displacement transduction circuits.

Figs. 19a, 19b and 19c are directed to magnetomotive reflection and bridge measurements. While the illustrated embodiment is directed to magnetomotive NEMS devices, it is to be understood that the spirit of the invention includes all types of NEMS  
15 devices regardless of the means of inducing motion, such as electrostatic, thermal noise, acoustic and the like. Fig. 19a is a schematic diagram illustrating the magnetomotive reflection where there is only one NEMS device producing a signal, and Fig. 19b is a schematic diagram illustrating bridge measurements where there are two NEMS devices producing signals balanced against each other. In both measurements,  
20 a network analyzer 68 or other oscillator supplies a drive voltage,  $V_{in}$ . In the bridge measurements in Fig. 19b,  $V_{in}$  is split into two out-of-phase components by a power splitter 70 before it is applied to ports 64 and 66.  $R_L$  is the input impedance and  $R_s$  the source impedance of the network analyzer 68. In the illustrated embodiment both  $R_s = R_L = 50\Omega$ .

25 The NEMS device 60b is modeled as a parallel RLC network in Fig. 19b, with a

complex mechanical impedance,  $Z_m(\omega)$  and a DC coupling resistance,  $R_e$ .  $\Delta R$  is the DC mismatch resistance between the NEMS devices 60a and 60b two arms of the bridge. The transmission lines, especially in bridge measurements at high frequencies, can disturb the overall phase balance if they have unequal electrical path lengths. Fig.

5 19c is a scanning electron microscope (SEM) micrograph of a representative bridge device of Fig. 19b, made out of an epitaxially grown wafer with 50nm-thick n+ GaAs and 100-nm-thick intrinsic GaAs structural layers on top of a 1  $\mu\text{m}$  thick AlGaAs sacrificial layer showing NEMS beams or devices 60a and 60b extending between detection port 62 and actuation ports 64 and 66. The Ohmic contact pads appear rough  
10 in the micrograph. The doubly clamped beams 60a, 60b have dimensions of 8  $\mu\text{m}$  ( $L$ ) x 150 nm ( $w$ ) x 500 nm ( $t$ ) and in-plane fundamental flexural mechanical resonance frequencies of about 35 MHz.

The balanced circuit shown in Fig. 19(b) with a NEMS transducer 60b on one side of the bridge and a matching effective resistor 60a of resistance,  $R \approx R_e$  on the  
15 other side, is designed to improve the detection efficiency. The voltage,  $V_o(\omega)$  at the read-out port 62 is nulled for  $\omega \neq \omega_0$  by applying two 180° out of phase voltages to the drive port 64 and drive port 66 in the circuit. We have found that the circuit can be balanced with exquisite sensitivity, by fabricating two identical doubly clamped beam transducers on either side of the balance point 62, instead of a transducer and a  
20 matching resistor.

A representative device with the equivalent drive ports 64 and 66 and balance or detection point 62 is shown in the SEM microphotograph of Fig. 19(c). In such devices, we almost always obtained two well-separated mechanical resonances, one from each beam transducer 60a, 60b, with  $|\omega_2 - \omega_1| \gg \omega_1 / Q_1$  where  $\omega_1$  and  $Q_1$  are the  
25 resonance frequency and the quality factor of resonance of the transducers 60a, 60b,

( $i=1,2$  respectively) as graphically depicted in Fig. 21. The graph of Fig. 21 indicates that in the vicinity of the either mechanical resonance, the system is well described by the mechanical transducer-matching resistor model of the operational circuit of Fig. 19(b). We attribute this behavior to the high  $Q$  factors ( $Q \geq 10^3$ ) and the extreme  
 5 sensitivity of the resonance frequencies to local variations of parameters during the fabrication process.

First, to clearly assess the improvements, we compare reflection and balanced bridge measurements of the fundamental flexural resonances of doubly clamped beams patterned from  $n^+$  (B-doped) Si as well as from  $n^+$  (Si-doped) GaAs. Electronic  
 10 detection of mechanical resonances of these types of NEMS transducers 60a, 60b without metallization layers have proven to be challenging, since for these systems the two-terminal impedances can be quite high;  $R_e \geq 2 \text{ k}\Omega$  and  $R_m \ll R_e$ . Magnetomotive beams normally require metallization in order to be driven, but in the case of bridge measurements the measurements are so sensitive that nonmetalized magnetomotive  
 15 semiconductors beams can be used. Nonetheless, with the bridge technique described here, we have detected fundamental flexural resonances in the  $10 \text{ MHz} < f_0 < 85 \text{ MHz}$  range for B-doped Si transducers and in the range from  $7 \text{ MHz} < f_0 < 35 \text{ MHz}$  range for Si-doped GaAs beams. In all our measurements, the paradigm that  $R_m \ll R_e$  remained true as  $R_m \leq 10 \text{ }\Omega$  and  $R_e$  remained in the range of  $2 \text{ k}\Omega < R_e < 20 \text{ k}\Omega$ .

20 Here, we focus on our results from  $n^+$  Si beams. These devices were fabricated from a B-doped Si on insulator wafer, with Si layer and buried oxide layer thicknesses of 350 nm and 400 nm, respectively. The doping was done at 950 °C and the average dopant concentration was estimated as  $N_a \approx 6 \times 10^{19} \text{ cm}^{-3}$  from the average sheet resistance,  $R_{\square} \approx 60 \text{ }\Omega$ , of the sample. The fabrication of the actual devices was  
 25 performed using conventional or standard methodologies employing optical lithography,

electron beam lithography and lift off steps followed by anisotropic electron cyclotron resonance (ECR) plasma and selective HF wet-etches. After fabrication, samples were glued into a chip carrier and electrical connections were provided by Al wire bonds. The electromechanical response of the bridge at the point 62 was measured in  
5 a magnetic field generated by a superconducting solenoid.

Fig. 20a is the graph of a doubly-clamped, B-doped Si beam resonating at 25.598 MHz with a  $Q$  about  $3 \times 10^4$  measured in reflection configuration in upper curves 72 and in bridge configuration for magnetic field strengths of  $B=0,2,4,6$  T in lower curve 74. The drive voltages are equal. The background is reduced by a factor of about 200 in  
10 the bridge measurements. The phase of the resonance in the bridge measurements are shifted  $180^\circ$  with respect to the drive signal as shown in Fig. 21. Fig. 20b is a graph of the amplitude of the broadband transfer functions for both configurations. The coupling between the actuation and detection ports in the bridge circuit is capacitive.

In particular, Fig. 20(a) shows the response of a device with dimensions  $15 \mu\text{m}$  (L) x 500 nm (w) x 350 nm (t) and with  $R_e \approx 2.14 \text{ k}\Omega$ , measured in the reflection (upper  
15 curves) 72 and bridge configurations for several magnetic field strengths in curves 74. The device has an in-plane flexural mechanical resonance at 25.598 MHz with a  $Q \approx 3 \times 10^4$  at  $T \approx 20 \text{ K}$ . The DC mismatch resistance,  $\Delta R$  was about  $10 \Omega$ . Note that a background reduction of a factor of about  $200 \approx R_e / \Delta R$  was obtained in the bridge  
20 measurements as shown in the analysis below.

Fig. 20 (b) is a graph which shows a measurement of the wideband transfer functions for both configurations for comparable drives at zero magnetic field. Notice the dynamic background reduction by a factor of at least 100 in the relevant frequency range.

25 In metallized SiC beams 60a, 60b with  $R_e$  about  $100 \Omega$  and embedded within the

bridge configuration, we were able to detect mechanical flexural resonances deep into the VHF band ( $R_m$  about  $1\Omega$ ). Fig. 21 is a graph which depicts a data trace of the in-plane flexural mechanical resonances of two  $2\text{ }\mu\text{m}$  ( $L$ ) x  $150\text{ nm}$  ( $w$ ) x  $80\text{ nm}$  ( $t$ ) doubly clamped SiC beams 60a, 60b. Two well-separated resonances are extremely prominent  
 5 at 198.00 and 199.45 MHz, respectively, with  $Q$  factors about  $10^3$  at  $T \approx 4.2\text{ K}$ . These beams were fabricated with Al and Ti top metallization layers of thicknesses of  $20\text{ nm}$  and  $3\text{ nm}$ , respectively, using a process described in below in regard to SiC beam fabrication.

NEMS devices 60a, 60b configured in a bridge can effectively be regarded as a  
 10 two-port device with isolated actuation-detection ports 64-62, 66-62. Obviously, the coupling between the two ports 64, 66 is not solely of a mechanical nature, but the mechanical response dominates the electromechanical transfer function due to the dynamical nulling of the electronic coupling between the ports 64, 66.

We have recently demonstrated continuous frequency tracking by a phase  
 15 locked loop (PLL) of the fundamental mechanical resonance a doubly clamped beam 60a or 60b configured in a bridge. Since the source impedance,  $R_s$  due to the power splitter is symmetric in both arms of the bridge, it is not explicitly incorporated into  $Z_{eq}'(\omega)$ , but can be regarded as part of  $R_e$ . In fact, replacing  $R_e$  with  $R_e + R_s$  would produce the more general form.

20 The voltage at point 62 in the circuit can be determined as

$$V(\omega) = -\frac{V_{in}(\omega)(\Delta R + Z(\omega))}{(Z_m(\omega) + \Delta R)(1 + \frac{R_e}{R_L}) + R_e(2 + \frac{R_e}{R_L})} = -\frac{V_{in}(\omega)}{Z_{eq}'(\omega)}(\Delta R + Z_m(\omega))$$

4. 3

by analogy to equation 4.1. At  $\omega = \omega_0$ , we can define a detection efficiency similar to

that of equation 4.2 for the signal  $S$  and background  $B$ :

$$\frac{S}{B} = \frac{R_m}{\Delta R} \quad 4.4$$

Given that  $\Delta R$  is small, the detection efficiency becomes considerably higher  
 5 than for the one port case. In the vicinity of the resonance, the background is  
 suppressed by a factor of order  $R_e / \Delta R$ , as confirmed by the measurements of Fig.  
 20(a). The inherent resistance mismatch  $\Delta R$ , due to the fabrication, however, is not the  
 ultimate limit to the background reduction.

Further balancing and hence background reductions can be obtained through  
 10 insertion of a variable attenuator 64a and a phase shifter 66a in the opposite arms. The  
 attenuator 64a will balance out the mismatch more precisely, while the phase shifter  
 66a will compensate for the phase imbalance created by the insertion of the attenuator  
 64a.

At higher frequencies, however, the circuit model of Fig. 19b, and hence the  
 15 above expressions, become imprecise as is evident from the measurements of the  
 transfer function. Capacitive coupling becomes dominant between the actuation ports  
 64 and 66 and the detection or balance port 62 at high frequencies as displayed in Fig.  
 20(b), and this acts to reduce the overall effectiveness of the technique. With careful  
 design of the circuit layout and the bonding pads, such problems can be minimized.

20 Even further signal improvements can be obtained by addressing the significant  
 impedance mismatch problem,  $R_e \gg R_L$ , which exists between the output impedance,  
 $R_e$  and the amplifier input impedance,  $R_L$ . In the illustrated embodiment, e.g. in the  
 measurements displayed in Fig. 20(a), this output impedance mismatch causes a signal  
 attenuation estimated to be of order about 40 dB. Output impedance matching circuits

62a can be used to avoid the mismatch between the beams and the load resistance.

### *Energy Dissipation in NEMS Devices*

Measurements on nanometer-scale doped beam transducers offer insight into  
5 energy dissipation mechanisms in NEMS devices, especially those arising from NEMS  
surfaces and surface adsorbates. In the frequency range investigated,  $10 \text{ MHz} < f_0 < 85$   
MHz, the measured  $Q$  factors of  $2.2 \times 10^4 < Q < 8 \times 10^4$  in B-doped Si beams is a factor of  
2 to 5 higher than those obtained from metallized beams. The comparison is strictly a  
qualitative one. We have compared  $Q$  factors of eight metallized and fourteen doped Si  
10 beams measured in different experimental runs, spanning the indicated frequency  
range. It has been suggested that both metallization layers and impurity dopants can  
make an appreciable contribution to the energy dissipation. Our measurements seem to  
confirm that at nanometer scales, metallization overlayers can significantly reduce  $Q$   
factor. Second, the high  $Q$  factors attained and the surfaces that are free of metal films  
15 make these doped beams excellent tools for the investigation of small energy  
dissipation changes due to surface adsorbates and defects. In fact, efficient *in situ*  
resistive heating in doped beams through  $R_e$  has been shown to facilitate thermal  
annealing and desorption of surface adsorbates thereby yielding even higher  $Q$  factors.  
These devices are promising for studies of adsorbate-mediated dissipation processes.  
20 In summary, we have developed a broadband, balanced radio frequency bridge  
technique for detection of small NEMS displacements. This technique may prove useful  
for other high frequency high impedance applications such as piezoresistive  
displacement detection. The technique, with its unique advantages, has enabled  
electronic measurements of mechanical resonances from systems that would otherwise  
25 be essentially unmeasurable.



### *Ultra High Frequency Silicon Carbide Nanomechanical Transducers*

Nanomechanical transducers with fundamental mode resonance frequencies in the ultra-high frequency (UHF) band are fabricated from monocrystalline silicon carbide thin film material, and measured by magneto motive transduction, combined with balanced bridge read out circuit. The highest frequency among units which have been fabricated prior to the invention is a measured resonance of 632 MHz. The technique described here also holds clear promise in accessing the microwave L-band frequencies of mechanical motion, which carries great hope in studying the physics of mechanical motion at the mesoscopic scale, as well as in developing brand new technologies for the next generation of nanoelectromechanical systems (NEMS).

In the illustrated embodiment, we disclose the fabrication and measurement of ultra-high frequency silicon carbide nanomechanical transducers. Our measurement, which is based on the magnetomotive transduction has successfully detected resonance with frequencies above 600 MHz. Further, it is easy to see that our technique is not limited to the already achieved UHF frequency range. Microwave L-band (1 to 2GHz) is also expected to be easily accessible by the same measurement setup with minor optimizations. The device fabrication process is similar to that described in Y. T. Yang et.al., Appl. Phys. Lett., 78, 162-164 (2001) with minor differences in terms of etch mask selection. The approach used here for nanometer-scale single crystal, 3C-SiC layers is not based upon wet chemical etching and/or wafer bonding. Especially noteworthy is that the final suspension step in the surface nanomachining process is performed by using a dry etch process. This avoids potential damage due to surface tension encountered in wet etch processes, and circumvents the need for critical point drying when defining large, mechanically compliant devices.

The starting material for device fabrication is a 259-nm-thick single crystalline 3C-SiC film heteroepitaxially grown on a 100 mm diameter (100) Si wafer. 3C-SiC epitaxy is performed in an RF induction-heated reactor using a two-step, carbonization-based atmospheric pressure chemical vapor deposition (APCVD) process. Silane and propane are used as process gases and hydrogen is used as the carrier gas. Epitaxial growth is performed at a susceptor temperature of about 1330 °C. 3C-SiC films grown using this process have a uniform (100) orientation across each wafer, as indicated by x-ray diffraction. Transmission electron microscopy and selective area diffraction analysis indicates that the films are single crystalline. The microstructure is typical of epitaxial 3C-SiC films grown on Si substrates, with the largest density of defects found near the SiC/Si interface, which decreases with increasing film thickness. A unique property of these films is that the 3C-SiC/Si interface is absent of voids, a characteristic not commonly reported for 3C-SiC films grown by APCVD.

Fabrication begins by defining large area contact pads by optical lithography. A 60-nm-thick layer of Cr is then evaporated and, subsequently, standard lift-off is carried out with acetone. Samples are then coated with a bilayer polymethylmethacrylate PMMA resist prior to patterning by electron beam lithography. After resist exposure and development, 30–60 nm of Cr is evaporated on the samples, followed by lift-off in acetone. The pattern in the Cr metal mask is then transferred to the 3C-SiC beneath it by anisotropic electron cyclotron resonance (ECR) plasma etching. We use a plasma of  $\text{NF}_3$ ,  $\text{O}_2$ , and Ar at a pressure of 3 mTorr with respective flow rates of 10, 5, 10 sccm, and a microwave power of 300 W. The acceleration DC bias is 250 V. The etch rate under these conditions is about 65 nm/min.

The vertically etched structures are then released by controlled local etching of the Si substrate using a selective isotropic ECR etch for Si. We use a plasma of  $\text{NF}_3$

and Ar at a pressure of 3 mTorr, both flowing at 25 sccm, with a microwave power 300 W, and a DC bias of 100 V. We find that  $\text{NF}_3$  and Ar alone do not etch SiC at a noticeable rate under these conditions. The horizontal and vertical etch rates of Si are about 300 nm/min. These consistent etch rates enable us to achieve a significant level  
5 of control of the undercut in the clamp area of the structures. The distance between the suspended structure and the substrate can be controlled to within 100 nm.

After the structures are suspended, the Cr etch mask is removed either by ECR etching in an Ar plasma or by a wet Cr photomask etchant (perchloric acid and ceric ammonium nitrate). The chemical stability and the mechanical robustness of the  
10 structures allow us to perform subsequent lithographic fabrication steps for the requisite metallization step for magnetomotive transduction on the *released* structures.

Suspended samples are again coated with bilayer PMMA and after an alignment step, patterned by electron beam lithography to define the desired electrodes. The electrode structures are completed by thermal evaporation of 5-nm thick Cr and 40-nm-thick Au  
15 films, followed by standard lift-off. Finally, another photolithography step, followed by evaporation of 5 nm Cr and 200 nm Au and conventional lift-off, is performed to define large contact pads for wire bonding.

SEM micrographs of a completed device are shown in Fig. 22. The photos Figs. 22a and 22b are the top view and sideview respectively, of the device region. The large  
20 area finger pads 76 are formed by thermally evaporated metal films of 6 nm of Cr for cohesion, followed by 80 nm of Au. The fine structures 78 of the device, defined by electron beam lithography, are covered by 36 nm of nickel film, deposited by electron beam evaporation. Such metal films, including Ni and Au, serve dual purposes, which are used as etch masks, and used for electrical conduction.

25 During the anisotropic electron cyclotron resonance (ECR) etch perpendicular to

the wafer surface, the metal films comprising part of structure 78 serve to protect the mono crystalline 3C-SiC thin film underneath them. This first etching step exposes the substrate silicon material in areas not covered by metals. The following second ECR etching step, which slowly removes silicon material isotropically, will suspend the metallized silicon carbide beams 78 from the substrate. Each device 10 is comprised of two nominally identical doubly clamped beams 78. Fig. 22c and 22d are zoom-in views of one of the two beams 78 in a device 10. Beam suspension can best be seen from the photo in Fig. 22d. Also from these photos, we can measure the geometry of the suspended beams to be roughly: a  $1.25\text{ }\mu\text{m}$  length  $l$ , a  $0.18\text{ }\mu\text{m}$  width  $w$ , and a  $0.075\text{ }\mu\text{m}$  thickness,  $t$ . The thickness for the SiC film is obtained by subtracting 36 nm of the nickel thickness from a measurement of the beam overall thickness or height, since the nickel thickness reduction during the entire etching process is calibrated as negligible.

The metal masks which are used are retained as the conducting layer needed for magnetomotive transduction. A typical beam with nickel metallization has a measured resistance of about 90 Ohms, with the resistance mismatch in between the two beams 78 in the same device to be within 1-2%.

The sample is subsequently mounted on the sample holder (not shown), and wire bonded to 50 Ohm microstrip lines (not shown), which in turn are coupled to 50 Ohm coax cables (not shown). The cables and connections linked to the device finger pads 76a and 76b in the bridge circuit in Fig. 23 are made nearly identical, reaching up to the two output connectors of the  $180^\circ$  power splitter 80, which divides the driving power from port 82 of the HP8720C network analyzer 84 into two equal partitions, but with a phase difference of  $180^\circ$ . In a cryogenic measurement the device 10 sits in a dipper or instrument column, whose vacuum can or sample chamber is evacuated and immersed into liquid helium. An uniform static magnetic field is applied by a

superconducting magnet (not shown), which has a field direction perpendicular to the doubly-clamped beams 78. When the RF current runs through the conducting layer of the beams 78, forces at the RF driving frequencies will be experienced by the beams 78. If the driving force frequency does not match the mechanical resonance frequency of beams 78, induced mechanical motion is minimal.

Terminal 86 will be the virtual ground in the ideal case, where the two beams 78 are exactly identical, as are the two branches of circuit components connected to them. Nonideality will introduce a residual background shift from the ideal virtual ground, as well as slightly different resonance frequencies of the two beams 78 in the device for the same mode. When the driving frequency matches the fundamental mode mechanical resonance frequency of one of the beams 78, resonant mechanical motion will occur for that beam 78. Such mechanical motion, which is perpendicular to the magnetic field, will induce an EMF voltage at the same frequency. This EMF voltage will act as an additional electrical generator, and affect the power transmitted out from terminal 86 of the device towards the detector port 88. Such power is then amplified and detected at port 88 of the network analyzer 84.

In the language of network analysis, we measure the frequency dependence of the forward transmission coefficient  $S_{21}$  of the network. As known from definition,  $|S_{21}|^2$  represents the power delivered to a matched load over the power incident on the input port. Information about the mechanical motion is revealed as resonance peaks in the spectra.

When the direction of applied magnetic field is in the plane of the wafer surface 90, which is the plane of Fig. 22a, and perpendicular to the beam 78, the direction of motion is perpendicular to the wafer surface 90 and is referred to as the out-of-plane resonance. A similar flexural mode, called in-plane resonance, will be excited when the



boundary conditions; for the fundamental mode of a doubly clamped beam  $\eta = 3.57$ .

Assuming the correction due to the electrode layer (layer 2) is small, we can define a correction factor  $K$ , to allow direct comparison with the expression for homogeneous beam

$$5 \quad f = \frac{\eta}{L^2} \left( \frac{E_1 I_{10}}{\rho_1 A_{10}} \right)^{\frac{1}{2}}$$

, where

$$K = \frac{E_1 I_1 + E_2 I_2}{E_1 I_{10}} \frac{1}{1 + \frac{\rho_2 A_2}{\rho_1 A_1}}$$

In this expression,  $I_{10}$  is the moment calculated in the absence of the second layer. The correction factor  $K$  can then be used to obtain a value for the *effective geometric factor*,  
 10  $[t / L^2]_{\text{eff}}$  for the measured frequency. Further nonlinear correction terms, of order higher than  $[t / L^2]_{\text{eff}}$  are expected to appear if the beams are under significant tensile or compressive stress. The linear trend of our data, however, indicates that internal stress corrections to the frequency are small.

The measured resonance frequencies are about 30% lower than such estimates.  
 15 The discrepancies are not surprising, comparing to what was encountered in our previous work at a lower frequency range. In particular, when the size of the device shrinks down, the role of surface, defects and non-ideal clamping etc. will become increasingly important. These factors are not considered in such predictions.

The in-plane resonance data is shown in Fig. 24, whereby magnetic field is 8  
 20 Tesla, driving power is -60dBm, with a resolution bandwidth equal to 10Hz. The frequency dependence of the forward transmission coefficient is plotted. The insert shows the projection of the complex function onto the  $S_{21}$  plane. Two resonance peaks

are observed at about  $180^\circ$  phase difference, as expected. In these data, information about both the mechanical transducer and the electrical connections is presented. To extract information about the mechanical resonant structure, we subtract the background, which is also a complex-valued function of frequency, fitted from data  
 5 points taken away from the resonance peaks. After subtracting background, the amplitude of the resulting function is plotted in Fig. 25. Within experimental error the de-embedded amplitude peaks can be fitted to Lorentzian shape and the peak height is roughly proportional to  $B^2$ , as expected.

The amplitude axis of Fig. 25 is normalized, so that its value represents the  
 10 signal voltage referred back to the input of the cryoamp 92. Such normalization can be easily done using the definition of network forward transmission coefficient, together with the knowledge of the gain (48dB) of the amplifier 92. In such estimates, we ignore the loss from the coax cables. Also ignored is the effect from the impedance mismatch at the device output, which in our case should only contribute a factor in the order of  
 15 unity. Under such simplification, the signal voltage referred to the input of the cryoamp 92 represented in Fig. 25, can be considered approximately the EMF voltage generated by the magnetomotive transduction, which can be expressed by

$$V_{emf} \sim B L 2\pi f_0 A \quad 5.1$$

20

where  $L$  is the length of the beam,  $f_0$  is the resonance frequency,  $B$  is the magnitude of the magnetic field, and  $A$  is the displacement amplitude of the mechanical motion. We thus obtain the maximum amplitude of the motion is about  $7 \times 10^{-3} \text{ \AA}$  under 8 Tesla magnetic field.

25 Using the same expression, we can also estimate the displacement sensitivity, if



the noise voltage per  $\sqrt{\text{Hz}}$  is known. In general, the detection sensitivity is limited by the Johnson noise from beam resistance, and the noise from the pre-amplifier 92. These two noise sources are comparable to each other, since the experiment is done at liquid helium temperature. Beam resistance is typically a few teens of Ohms, and the noise temperature of the MITEQ cryogenic amplifier 92 is in the order of a few Kelvin in the frequency range of interest. The combined noise is effectively a noise temperature of about 10 K referred to the input, which corresponds to a noise voltage per  $\sqrt{\text{Hz}}$  of  $150 \text{ pV}/\sqrt{\text{Hz}}$ . This in turn gives a displacement sensitivity of about  $5 \times 10^{-5} \text{ \AA} / \sqrt{\text{Hz}}$ . In reality, the noise estimated from Fig. 25 is higher than the above values by a factor of a \_\_\_\_.

10 This additional noise reflects the receiver sensitivity of the network analyzer 84.

For the purpose of this particular illustration, we did not attempt to optimize the noise performance of the system. However, it is essentially trivial to do so, by adding a reasonably low-noise second stage amplifier (not shown) with -40dB gain after the cryoamp 92, so as to utilize the full capabilities of the ultra low noise feature of the cryoamp 92.

As a first order approximation, we know that the resonance frequency for in-plane case

$$f_0 = \sqrt{\frac{E}{\rho}} \frac{W}{L^2} \quad 5.2$$

20 Where W, L is the width and length of the beam, respectively. E is the Young's modulus,  $\rho$  is the mass density. Combining Eqn. 5.1 and 5.2 we have

$$V_{emf} \propto B A \sqrt{\frac{E}{\rho}} \frac{W}{L} \quad 5.3$$

From Eqn. 5.2, we know that shrinking the size of the beam in all three dimensions by the same ratio, using the device described above as the starting point, can easily get the resonance frequency into the microwave L-band. Such scaling down is readily achievable by current technology of e-beam lithography. On the other hand, Eqn. 5.3 tells us that the signal amplitude will not be significantly reduced, as long as we keep the same B field, the same material, and similar amplitude of mechanical motion.

In conclusion, we have demonstrated the measurement of silicon carbide nanomechanical transducers with fundamental resonance frequencies in the UHF range, and the microwave L-band frequencies as well by the same technique. This gives access to frequency bands of mechanical motion, which were never inaccessible before.

#### 15 *Frequency Tuning of MEMSINEMS Transducers by the Lorentz Force*

The resonance frequency of a magnetomotive NEMS transducer can be fine tuned by varying the static stress applied to the resonating beam by means of a Lorentz force device from a DC current passed through the beam. We have performed all of our measurements on doubly clamped beams 94 such as those displayed in the SEM photograph of Fig. 26. These beams were microfabricated out of GaAs and Si. In order to electrically couple to these mechanical structures, we patterned thin Au or Al electrode layers of  $d \approx 50$  nm on top of the beams 94. Several beams 94 with different lengths of  $50 \mu\text{m} < L < 70 \mu\text{m}$  and with fixed  $w = 1.5 \mu\text{m}$  and  $t = 0.8 \mu\text{m}$  were used for the force tuning experiments, covering a frequency range of  $1 \text{ MHz} < \omega_0/2\pi < 3.5 \text{ MHz}$ . For investigations of the temperature variations of the frequency, several beams with

differing aspect ratios ( $4 \text{ MHz} < f < 40 \text{ MHz}$ ) were fabricated on the same chip, and the resonance frequency was recorded as the temperature was varied.

A magnetomotive excitation and detection scheme was used for the measurements. Briefly, a network analyzer 96 was used to drive an alternating current (AC) along an electrode (not shown) on top of the beam 94, which was placed in the bore of a superconducting magnet (not shown) at 4.2 K. The Lorentz force due to this AC current excited the beams 94 and the electromotive force generated by the motion was detected by the network analyzer 94. The frequency shift data were obtained from the resonance curves by inspection.

The tuning force was introduced by running a direct current (DC) as well as the AC drive current through the electrode. The DC current in a constant magnetic field subjects the beam to a constant Lorentz force per unit length,  $\tau = IB$ , where  $I$  is the current and  $B$  the magnetic field. Two different geometries were investigated in these experiments. In the first case, the beams 94 were excited perpendicular to the plane of the chip (defined as the  $z$ -direction) and were subjected to a constant force in the same direction via the DC current flow. In the second case, the beam was rotated through a  $90^\circ$  angle with respect to the magnetic field and the excitation and the tensile forces were in plane ( $x$ - $y$  plane).

The motion of the doubly clamped beams 94 can be modeled by the beam equation:

$$\frac{\partial^4 u(x,t)}{\partial x^4} - \frac{\sigma A}{EI} \frac{\partial^2 u(x,t)}{\partial x^2} = -\frac{\rho A}{EI} \frac{\partial^2 u(x,t)}{\partial t^2}$$

where  $\sigma$  is the tensile or the compressive stress in the beam,  $A$  and  $I$  are the cross-sectional area and moment of the area, respectively.  $E$  is the Young's modulus and  $\rho$  is the mass density of the material as usual,  $t$  is time,  $x$  is the distance along the

beam and  $u$  is the displacement of the beam in the direction of excitation.

To present a more general discussion, we have included an internal stress term in the beam equation. However, our analysis below shows that internal stresses do not modify significantly the observed beam resonances. The frequency of the fundamental

5 resonance can be derived from the above equation as

$$\omega_0 = 1.03 \frac{t}{L^2} \sqrt{\frac{E}{12\rho} \left( 1 + \frac{12L^2\sigma}{4\pi^2 Et^2} \right)}$$

where  $t$  and  $L$  are the beam thickness and length respectively and  $E$  is the Young's modulus.

We have measured the resonance frequencies of up to 30 Si and GaAs beams.

10 Fig. 27 shows the measured fundamental frequency of the beams 94 as a function of the aspect ratio,  $t/L^2$ . The fact that we obtain linear dependence of the resonance frequency,  $f$  on  $t/L^2$  suggests that the corrections to  $f$ , due to various internal stresses in the beam 94 are very small. The measured  $E/\rho$  values from the slopes in Fig. 27 are only within 75 % of the calculated values. This, however, can be explained by the

15 frequency lowering effects of unintended undercuts in the semiconductor sacrificial layer that might change the effective length by up to 10 % and mass loading effects due to the electrode layers disposed on the beam for the magnetomotive current (not shown).

#### 20 *i) Lorentz Force Tuning:*

In Fig. 28, we present the Lorentz force tuning curves of the out-of-plane resonance of a 1.177 MHz beam 94. The frequency shift,  $\delta f_z/f_z$ , where  $\delta f_z$  is the change in frequency for a z-direction or out-of-plane excitation and  $f_z$  is frequency for a z-direction or out-of-plane excitation frequency, as a function of the applied DC current is

plotted at three different magnetic fields. The fact that the plots collapse onto the same curves shown in Fig. 30 reassures that this effect is indeed a force tuning effect. The apparent curvature at the lowest fields is due to the heating effect of the DC current, as will be discussed below. We note that qualitatively similar curves were obtained for four  
 5 different GaAs samples with  $1 < f < 3$  MHz.

Fig. 29 shows the normalized in plane frequency shift,  $\delta f_{xy}/f_{xy}$ , of the same beam 94 for in-plane excitation as a function of the current for different magnetic field strengths. The lack of symmetry in the data becomes more evident as the magnetic field strength is increased. The tuning plotted as a function of the applied force per unit  
 10 length in Fig. 31 implies that the force tuning effects in this plane are very weak and are probably obscured by the frequency lowering effects of heating.

#### *li) Thermal Tuning:*

The temperature variations of the normalized out-of-plane and in-plane  
 15 frequencies of a GaAs beam covered with a thin Au layer are shown in Fig. 32. It is important to note that the two modes exhibit different temperature coefficients with the stiffer mode showing the least change. In this case the beam dimensions  $w \times t \times L$  were  $1.5 \times 0.8 \times 70$  microns. The out-of-plane and in-plane resonances were at  $f_z = 1.177$  MHz and  $f_{xy} = 1.838$  MHz respectively. A similar effect was observed in a Si beam with  
 20 slightly higher frequencies ( $f_z = 2.830$ ,  $f_{xy} = 2.328$  MHz).

The data of Fig. 32 is suggestive that thermal tuning will be weak in very stiff structures. This expectation is verified by measuring temperature dependence of the resonance frequencies of a number of beams 94 with a range of frequencies. The data for Si and GaAs are displayed in Figs. 33 and 34 respectively. Also plotted on the data  
 25 for both materials is the variation of the sound velocity given that the density change is

negligible over the temperature range. Any conventional source of heating and cooling can be employed to vary the temperature.

*i) Lorentz Force*

5 We have argued above that the intrinsic stresses in the beams do not contribute appreciably to the observed resonance frequencies of our structures (see Fig. 27). We will analyze the tuning problem by assuming a neutral beam and by adding a stress term due to the constant Lorentz force. We therefore start by reviewing the response of a clamped beam to constant stresses axial with the beam, which we later relate to the  
10 Lorentz Force. The equation of motion for small amplitudes around the equilibrium point, is

$$\frac{\partial^4 u(x,t)}{\partial x^4} - \frac{\sigma A}{EI} \frac{\partial^2 u(x,t)}{\partial x^2} = -\frac{\rho A}{EI} \frac{\partial^2 u(x,t)}{\partial t^2} \quad 6.1$$

where  $\sigma$  is the tensile or the compressive stress in the beam,  $A$  and  $I$  are the  
15 cross-sectional area and moment of the area, respectively.  $E$  is the Young's modulus and  $\rho$  is the mass density of the material as usual. The frequency of oscillation for the stressed case,  $(\omega_0')$  for clamped boundary conditions, can be obtained by solving the above equation:

$$20 \quad \omega_0' = 1.03 \frac{1}{L^2} \sqrt{\frac{EI}{\rho A} \left(1 + \frac{AL^2 \sigma}{4\pi^2 EI}\right)} = \omega_0 \sqrt{1 + \frac{3L^2 \sigma}{\pi^2 Et^2}} \quad 6.2$$

In this equation,  $L$  and  $t$  are the length and thickness of the beam 94, respectively. The resonance frequency can increase or decrease depending on the

nature of the stress, i.e. compressive or tensile.

A small constant transverse force per unit length modifies the equilibrium shape of the beam 94. A beam 94 under the effect of such a pull assumes elastically a shape described by

5

$$u(x) = \frac{\tau}{24EI} x^2 (x-L)^2 \quad 6.3$$

where  $\tau$  is the constant force per unit length on the beam. This force causes the beam to elongate and hence results in a tensile stress. The tensile stress due to  $\tau$  is

10 given by

$$\sigma = E \frac{\Delta L}{L} = \frac{1}{60480} \frac{L^6}{Et^2} \tau^2 = \frac{1}{420} \frac{L^6}{Ew^2t^6} \tau^2 \quad 6.4$$

Therefore the new resonant frequency using eqn. 6.2 is

$$\omega_0' = \omega_0 \sqrt{1 + \frac{1}{140\pi^2} \left(\frac{L}{t}\right)^8 \frac{\tau^2}{E^2 w^2}} \quad 6.6.$$

15

Note that the frequency shift is positive for all transverse forces.

The expression for frequency shift for the constant force per unit length,  $\tau = IB$ , takes the form

$$\omega_0' = \omega_0 \sqrt{1 + \frac{1}{140\pi^2} \left(\frac{L}{t}\right)^8 \frac{I^2 B^2}{E^2 w^2}} \quad 6.7$$

20

The prefactor for our GaAs beams with  $L/t \approx 50$  and  $w = 1 \mu\text{m}$  is of order  $10^0$  in

metric units. The maximum force per unit length applied is  $4 \times 10^{-3} \text{ N/m}$ . Therefore we can safely expand the frequency shift as

$$\Delta\omega \approx \omega_0 \left( \frac{1}{280\pi^2} \left( \frac{L}{t} \right)^8 \frac{I^2 B^2}{E^2 w^2} \right) \quad 6.8$$

This expression estimates a normalized frequency shift of order  $10^{-5}$  -  $10^{-6}$  for our  
5 beams 94. Our measurements, however, deviate from the above expressions in several  
significant ways. First, the measured frequency shift is asymmetric for the resonance in  
the z direction, and we encounter with negative frequency shift for forces pulling the  
beam towards the substrate 95. The effect we are observing is significantly larger and  
is linear in both variables  $B$  and  $I$ .

10           The method to apply the constant force per unit length, however, causes complications in the case of Lorentz force tuning. The constant current  $I$  produces a local temperature increase estimated to be about 5 to 10 K. Therefore the measured frequency shift is a more detailed function of the applied current and hence force:

$$\Delta f = \Delta f_{\text{tuning}}(I, B) + \Delta f_{\text{heating}}(I)$$

This effect becomes more apparent as the magnetic field strength  $B$  goes to zero. In the case of  $B=0$ , we would expect a completely symmetric curve in  $I$ . As seen in Fig. 29 the frequency shift curve becomes more symmetric as the tuning force becomes smaller. In Fig. 35 we plot the data in Fig. 29 after subtracting the even component, which we assume is due to heating. Note the similarity of Fig. 35 to Fig. 28. However, the effect in Fig. 35 is an order of magnitude smaller.

We do not understand the origin of the asymmetric tuning observed in both cases. Asymmetric tuning of this kind could be observed in buckled structures, however



observed beam resonance frequencies in our experiments indicate that our beams are far away from the buckling transition. The interesting temperature dependencies of the resonance frequencies in Figs. 32 and 33 suggest that resonant frequency shift with temperature is not responsible for the observed behavior. The observed effects might  
5 be due the stresses formed in the semiconductor-contact metal bilayer. Single component beams made out of polycrystalline metals and single crystalline highly doped semiconductors eliminate the above mentioned stresses.

### *Ultimate Limits Of Displacement Detection With Flexural And Torsional Transducers*

#### 10 *Using Magnetomotive Transduction*

In the illustrated embodiment we quantify the performance of the magnetomotive detection technique in the context of micromechanical transducers. We outline the factors which limit its displacement sensitivity at frequencies from 1 MHz to 1 GHz. We evaluate the sensitivity for realistic systems and instruments, and show that it is  
15 possible to attain the thermomechanical noise limit of sensitivity at 1 GHz.

#### *i) Magnetomotive Transduction*

In the presence of a magnetic field, mechanical motion perpendicular to the field induces an electromotive force (EMF) perpendicular to both. An electrode on the  
20 moving object transmits the induced voltage signal or EMF to a detector. Let us evaluate the magnetomotive transduction of a mechanical transducer's motion. At frequencies close to a normal mode of the transducer, and at low amplitudes, its motion is well-described by a damped simple harmonic oscillator, with an effective mass  $m$  and effective spring constant  $k$ :

$$m \frac{d^2 y(t)}{dt^2} + \gamma m \frac{dy(t)}{dt} + ky(t) = F(t)$$

5  $\gamma$  denotes the damping coefficient arising due to the coupling of motion to internal and external degrees of freedom which cause dissipation. The value of  $m$  depends on the mode shape, and the value of  $k$  depends on how the force  $F$  is applied and the location at which the displacement  $z$  is measured. For a straight doubly-clamped beam of length  $L$ , thickness  $t$ , and width  $w$ , vibrating in its fundamental flexural mode in the  $t$  direction, the spring constant measured at the center of the beam for a uniform force is:

$$k = 32 E (t/L)^3 w$$

10 where  $E$  is the elastic constant of the material.

The EMF per unit length induced in the detection electrode at the coordinate  $x$  along the electrode is

$$V_0(x,t) = B \frac{dy(c,t)}{dt} \sin \theta(x)$$

15 where  $y$  is measured perpendicular to the field and  $\theta$  is the angle between the electrode and the field,  $B$ . By integrating along the length  $L_e$  of the detection electrode, the total voltage can be expressed in terms of a geometric factor  $\xi$  as:

$$\bar{V}_0(t) = \xi L_e B \frac{dy(t)}{dt}$$

20 For the fundamental flexural mode of a straight doubly-clamped beam,  $\xi = 0.53$  if displacement is measured at the center. At the resonance frequency  $\omega_0$ , then, the efficiency of magnetomotive transduction is given by:

$$V_0 \cong 2\pi\xi LBf_0 y_0$$

We thus define the responsivity  $R$  of the device as follows:

$$R = \frac{V_0}{y_0} = 2\pi\xi LBf_0$$

## 5 (ii) *Magnetomotive Circuit Model*

For a high  $Q$  transducer driven by the Lorentz force  $F = Bl_d L_d$  on a drive electrode of length  $L_d$ , magnetomotive transduction yields a Lorentzian line shape centered at the resonance frequency  $\omega_0 = \sqrt{k/m}$ :

$$V_0(\omega) = \frac{i\omega\omega_0^2 \xi L_e LB^2 / k}{\omega_0^2 - \omega^2 + i\gamma\omega} I_d(\omega)$$

10

For a straight doubly-clamped beam, the fundamental frequency for vibration in the thickness direction is:

$$f_0 = 1.03 \frac{t}{L^2} \sqrt{\frac{E}{\rho}}$$

where  $\rho$  is the mass density of the material.

15 The equation of motion has the same form as that of the voltage generated in a parallel LCR circuit schematically shown in Fig. 36, so analogous electrical parameters for the mechanical system can thus be defined as follows.

$$R_m = \frac{\xi L^2 B^2 \omega_0}{k} Q$$

$$C_m = \frac{k}{\xi L^2 B^2 \omega_0}$$

$$L_m = \frac{\xi L^2 B^2}{k}$$

The quality factor  $Q$  describes the dissipation of the energy of motion, and is related to the damping coefficient:  $\gamma = \omega_0/Q$ . Mechanical dissipation is thus represented by a mechanical resistance. For the fundamental resonance  $\omega_0$  of a doubly-clamped  
5 silicon beam,

$$R_m = .004444 \frac{B^2 Q}{wt^{\frac{1}{2}} f^{\frac{3}{2}}}$$

The amplitude of motion is proportional to the electrical amplitude across the LCR transducer by the responsivity.

In principle, the technique used to generate the transducer's motion is not directly  
10 relevant to its detection. However, in practice, due to space constraints on submicron transducers it is often convenient to use a single electrode on the surface of the beam for both drive and detection. In the magnetomotive scheme, by passing an alternating current through the electrode perpendicular to the field, an oscillating Lorentz force can be applied to the device. Our analysis will be divided into two qualitatively different  
15 cases: the one-port case, in which a single electrode serves as both magnetomotive drive and detection, and the two-port case, in which the detection electrode is separate. The two-port case is relevant to the measurement of the transducer's response to an external stimulus in the absence of magnetomotive drive.

20 *One-port case:*

The one-port circuit model is shown in the schematic of Fig. 37. The resistance

96,  $R_e$ , denotes the DC resistance of the electrode and the resistance 98,  $R_L$ , the detector's input impedance. The resistance 100,  $R_0$ , provides a large embedding impedance to make the drive a current source 102. The device is connected to the drive 102 by a  $50\Omega$  transmission line 104. The RLC circuit of Fig. 36 is coupled  
 5 between resistance 96,  $R_e$ , and ground.

*Two-port case:*

In the two port case with magnetomotive drive, the drive circuit is identical to the one port circuit of Fig. 37. The detection circuit as shown in Fig. 39 is completely  
 10 separate, except for a small reactive coupling. The detect electrode can be modeled as an ideal AC voltage source in series with the electrode resistance. The AC source voltage  $V'$  is proportional to the voltage across the RLC parallel circuit, or the motion of the transducer. The flow of current  $I_m$  in the measurement circuit affects the drive circuit by adding to the damping force in the equation of motion:

$$15 \quad \gamma \rightarrow \gamma' = \gamma + \kappa B L_e I_m / m$$

where  $L_e$  is the length of the detect electrode, and  $\kappa$  is a geometric factor to account for the two electrodes being at different locations on the structure. In the case of a straight beam with two identical parallel electrodes,  $L' = L$  and  $\kappa = 1$ . The circuit for the mechanical resonance is modified by the addition of a parallel resistance:

20

$$R_d = \frac{\xi B L^2}{\kappa L' I_m}$$

This approximation is valid near the resonance peak.

### *Coupling to the measurement circuit*

The most significant obstacle to magnetomotive detection at high frequencies is the efficient coupling of the transduced signal to the detector. As the device frequency is increased and its overall size is reduced, the dimensions of the detection electrode must be reduced proportionally, in order that the mechanical properties of the device are not ultimately dominated by electrode itself. Since the electrode's resistance scales as  $L/wt$ , it must be taken into account. For typical nanomechanical devices operating at 100 MHz and above, this source impedance  $R_s$  is much higher than the load impedance  $R_L$  of the detection circuit. If no attention is paid to the coupling circuit, the voltage measured by the detector can be substantially reduced.

### *One port case:*

In the one-port case, the most straightforward coupling option is to connect a detector to the device either directly or through a transmission line. If a standard RF amplifier with  $R_L = 50\Omega$  is used, then a transmission line of length  $\lambda/2$  acts as a 1-1 transformer, and we can substitute the equivalent circuit shown in Fig. 38. In this circuit configuration, the electrical response is not directly proportional to the motion of the transducer. For this reason it is appropriate to define the coupling efficiency as the ratio between the difference  $V_m$  in voltage at the detector on and off resonance, and the voltage  $V_0$  induced by the motion. On resonance, the mechanical part of the response is given by  $R_m$ , while off-resonance, it is essentially zero. Thus the coupling efficiency  $\varepsilon_1$  is given by:

$$\varepsilon_1 = \frac{V_m}{V_0} = \frac{V_{onresonance} - V_{offresonance}}{V_0} = \frac{(R_e + R_m)R_L^2}{(R_e + R_L)(R_e + R_m + R_L)^2}$$

Note that the coupling efficiency is reduced when the electrode resistance is large, and also when the mechanical resistance, or the transducer's responsivity, is large. The coupling can be improved by using a high-impedance detector such as a metal semiconductor field effect transistor (MESFET) (not shown), but the improvement is

5 only substantial if it is connected directly across the device.

*Two port case:*

In the two-port case, the most practical coupling strategy is to transform the source impedance down to the 50Ω input impedance of a standard low-noise RF amplifier.

10 Here we consider the simplest impedance transformation, a two-element L-section, as shown in Fig. 39.

The optimum choice for the reactive elements is:

$$L = \frac{R_S}{\omega} \sqrt{\frac{R_L}{R_S - R_L}}$$

$$C = \frac{1}{\omega R_S} \sqrt{\frac{R_S - R_L}{R_L}}$$

15 where  $R_S$  refers to the resistance of the detect electrode.

The measured voltage is then reduced by the factor:

$$\varepsilon_2 = \frac{V_m}{V_0} = \left( \frac{R_L}{R_S} \right)^{\frac{1}{2}} \left( \frac{R_S - R_L}{4R_S - 3R_L} \right)^{\frac{1}{2}}$$

For example, a signal from a 1kΩ electrode transformed to 50Ω is coupled with an efficiency of 0.11, compared to 0.0023 in the one-port case. It is clear that the two-port

20 configuration is preferable as long as there is sufficient space on the device for two

electrodes, and especially when the intent is to measure the response of the device to an external stimulus

*Parasitic reactance:*

- 5 At frequencies above 100 MHz, effect of parasitic reactance on the coupling circuitry must be considered. For a straight doubly clamped silicon beam transducer of length 3  $\mu\text{m}$ , width 200 nm and thickness 100 nm, vibrating at 100 MHz, the self-inductance for a 70 nm-wide electrode is negligible, at  $\sim 2 \text{ m}\Omega$ . The mutual impedance between two electrodes of width 70 nm, separated by 60 nm on the same transducer is  $\sim 1 \text{ m}\Omega$ . Their
- 10 capacitance is also negligible, at  $\sim 1 \text{ fF}$ . To a first approximation, the capacitance and inductance from these elements, to a first approximation, scale as  $L \log(L/w)$ , they are not expected to be important for standard geometries, well into the GHz frequency range. The most significant parasitic element is the capacitance between the ground plane on which the substrate rests, and the leads connecting the device to the
- 15 transmission line. For typical leads of width 100  $\mu\text{m}$  and length 500  $\mu\text{m}$ , on a silicon substrate of thickness 500  $\mu\text{m}$ , the shunt capacitance is  $\sim 150 \text{ fF}$ , or  $1 \text{ k}\Omega$  at 1 GHz. Since this capacitance shunts a detection electrode of similar impedance, it will reduce the coupling efficiency and ultimately the sensitivity of the measurement. In order to ensure efficient coupling at frequencies above 1 GHz, care must be taken to either
- 20 minimize the lead length, or to provide a proper transmission line to the device by fabricating a coplanar waveguide on the substrate.

*Sensitivity Analysis*

*System constraints:*

- 25 The sensitivity limit of the magnetomotive detection technique is a function of each



of the three components of the measurement: transduction, coupling, and amplification. As shown above, the transduction efficiency or responsivity, depends in a straightforward way on the physical dimensions of the device and the frequency of operation. The coupling efficiency of the readout circuit has the most potential for  
 5 optimization, as it depends on many parameters, including the finite resistance of the detection electrode, stray reactance, and the coupling circuit elements themselves. The input noise of the readout amplifier is taken to be fixed. In principle there are three ultimate noise sources for the measurement: the amplifier noise  $S_v^a$ , the Johnson noise  $S_v^J$  in the detection electrode, and the intrinsic thermomechanical vibration of the  
 10 transducer. The spectral density  $S_x^m$  of noise introduced by the measurement can be converted to the motion of the device as follows:

$$S_{X(1,2)}^M = \frac{1}{R^2} \left( S_v^J + \frac{1}{\mathcal{E}_{(1,2)}^2} S_v^a \right)$$

Our calculation demonstrates that the device and readout can be designed to reduce the contribution of the amplifier noise below the expected thermomechanical  
 15 noise, at frequencies up to 1 GHz.

To constrain the scope of the problem, we apply the general relationships developed in the analysis to the simple case of a straight doubly-clamped beam with one or two gold electrodes on its surface, vibrating in its fundamental normal mode. We further require that the device thickness be no less than 50 nm, and the drive and  
 20 detection electrodes be significantly thinner. Many applications have the additional requirement that the measurement circuit have negligible influence on the motion to be measured. In magnetomotive detection, the back-action or the perturbative effect of measurement is proportional to the current drawn by the measurement circuit.

*Transduction geometry:*

The geometry of a nanomechanical device is typically constrained by the thickness of the structural layer from which it is fabricated, or by the aspect ratio appropriate to the fabrication process or the application. For the simple flexural and torsional transducers shown below, there are only two independent parameters among ( $L$ ,  $t$ ,  $f_0$ ). Since we are particularly interested in high-frequency applications, we will calculate the geometry-related parameters of magnetomotive transduction in terms of ( $t$ ,  $f_0$ ) and ( $L/t$ ,  $f_0$ ).

Tables 1 and 2 show the frequency and responsivity in silicon for these two simple geometries. Table 1 lists the geometry-related parameters for flexural and torsional transducers. The force constant is measured at the beam's center 202 in the flexural case and at the edge of the paddle 200 in the torsional case as diagrammatically shown in Fig. 40. All numerical quantities have SI units. Table 2 lists the geometry-related parameters for typical flexural and torsional transducers.

TABLE 1

	fundamental flexure	torsion
frequency	$f_0 = 8800 \frac{t}{L^2}$	$f_0 = 1900 \frac{t^{3/2}}{L^{5/2}}$
force constant	$k = (6.60 \times 10^6) \omega t^{3/2} f_0^{3/2}$ or $k = (4.78 \times 10^{16}) \left( \frac{t}{L} \right)^5 f_0^{-1}$	$k = (2.58 \times 10^7) t^{11/5} f_0^{6/5}$ or $k = (4.22 \times 10^{14}) \left( \frac{t}{L} \right)^{11/2} f_0^{-1}$
length scale	$L = 95 (t / f_0)^{1/2}$	$L = 20.5 t^{3/5} f_0^{-2/5}$
responsivity	$R(\text{V/m}) = 312 (f_0 t)^{1/2} B$ or $R(\text{V/m}) = 29300 (t/L) B$	$R(\text{V/m}) = 129 f_0^{3/5} t^{3/5} B$ or $R(\text{V/m}) = 12000 \left( \frac{t}{L} \right)^{3/2} B$

TABLE 2

	flexure	torsion	flexure	torsion
frequency	100 MHz		1 GHz	
thickness	100 nm		50 nm	
length scale	3 $\mu\text{m}$	800 nm	670 nm	215 nm
force constant	21 N/m	41 N/m	115 N/m	140 N/m
responsivity, 8T	790 nV/ $\Rightarrow$	510 nV/ $\Rightarrow$	1.12 $\mu\text{V}/\Rightarrow$	1.08 $\mu\text{V}/\Rightarrow$

5 Doubly clamped beams 202 and torsional transducers 200 offer comparable magnetomotive responsivity in the RF frequency range. While their force constants and responsivities are similar, straight beams offer a distinct advantage over torsion paddle transducers. To achieve frequencies approaching 1 GHz with thickness no less than 50 nm, the torsional transducer must have torsion rods 204 with very low aspect ratio. For

10 example, for the 1 GHz transducer described in the table, this aspect ratio is 4. Not only is the structure difficult to fabricate, but the nonlinear coefficient in the restoring torque

is strong for torsion rods with such a small aspect ratio. This severely limits the linear dynamic range of any device application.

*Coupling:*

- 5           The coupling efficiency is governed by two contradictory requirements. The source impedance should be small, while at the same time the detect electrode should be small, in order to minimize mass-loading and possible damping effects. To simplify the analysis, we set an upper limit A on the ratio of electrode thickness to device thickness, which in principle would depend on the specific application. In the calculation  
10 we assume the electrode is optimal, having as large a cross-section as possible. For a straight beam, the resistance of the electrode is then given by:

$$R_s = \frac{L}{\lambda \sigma t w_e}$$

where  $\sigma$  is the conductivity of the electrode,  $\lambda$  is the wavelength of the driving signal,  $t$  is the beam thickness,  $L$  is the beam length and  $w_e$  is its width.

- 15 In the one-port case, the insertion loss or the ratio, power out/power in, of the coupling circuit is:

$$\varepsilon_1 = \frac{\alpha^2}{(1+\alpha)(1+\alpha + \frac{R_m \lambda \sigma t w}{L})}$$

where  $\alpha = \frac{\lambda \sigma R_L t w}{L}$ .

- 20 For typical devices with large aspect ratio  $L/t$ ,  $\alpha \ll 1$ , so the coupling efficiency can be approximated as  $\alpha^2$ .

In the two-port case, the insertion loss of the coupling circuit is:

$$\varepsilon_2 = \frac{1}{2} \alpha^{\frac{1}{2}} \left( \frac{1-\alpha}{1-.75\alpha} \right)^{\frac{1}{2}}$$

For typical devices with large aspect ratio  $L/t$ ,  $\alpha \ll 1$ , so the term in parentheses  
 5 can be neglected. Since parasitic reactances are small, this result is qualitatively valid  
 whether or not magnetomotive drive is used. If magnetomotive drive is used, however,  
 the presence of two electrodes on the beam requires  $x$  to be reduced by a factor of  
 approximately  $\sqrt{3}$ .

The coupling efficiency can be expressed in terms of the thickness or aspect ratio  
 10 of a straight beam:

$$\varepsilon_2 = 0.496(\sigma\lambda)^{\frac{1}{2}} \left( \frac{\rho}{E} \right)^{\frac{1}{8}} f_0^{\frac{1}{4}} t^{\frac{1}{4}} w^{\frac{1}{2}} \left( \frac{1-\alpha}{1-0.75\alpha} \right)^{\frac{1}{2}}$$

where 
$$\alpha = 0.99 R_L \sigma \lambda \left( \frac{\rho}{E} \right)^{\frac{1}{4}} f_0^{\frac{1}{2}} t^{\frac{1}{2}} w$$

$$\varepsilon_2 = 0.507(\sigma\lambda)^{\frac{1}{2}} \left( \frac{E}{\rho} \right)^{\frac{1}{4}} \left( \frac{t}{L} \right)^{\frac{3}{2}} f_0^{-\frac{1}{2}} \left( \frac{1-\alpha}{1-0.75\alpha} \right)^{\frac{1}{2}} \quad \text{for } w = t$$

where 
$$\alpha = 1.03 R_L \sigma \lambda \left( \frac{E}{\rho} \right)^{\frac{1}{2}} \left( \frac{t}{L} \right)^3 f_0^{-1} \quad \text{for } w = t$$

15

*Measurement sensitivity:*

The sensitivity of magnetomotive detection is limited by two sources of electrical noise: Johnson noise in the detection electrode itself, and noise at the amplifier input.

The spectral density for the detection electrode is:

$$\sqrt{S_V^J} = \sqrt{\frac{4k_B T L_e}{\lambda \sigma t w_e}}$$

- 5 This expression can be written in terms of the thickness or the aspect ratio of a straight beam:

$$\sqrt{S_V^J} = 2.01 \left( \frac{k_B T}{\sigma \lambda} \right)^{\frac{1}{2}} \left( \frac{E}{\rho} \right)^{\frac{1}{8}} f_0^{-\frac{1}{4}} t^{-\frac{1}{4}} w^{-\frac{1}{2}}$$

$$\sqrt{S_V^J} = 1.97 \left( \frac{k_B T}{\sigma \lambda} \right)^{\frac{1}{2}} \left( \frac{\rho}{E} \right)^{\frac{1}{4}} \left( \frac{L}{t} \right)^{\frac{3}{2}} f_0^{\frac{1}{2}} \quad \text{for } w = t$$

- 10 Combining the responsivity, the coupling efficiency, and the electrical noise sources, we obtain the spectral displacement sensitivity of the 2-port measurement on a straight doubly-clamped beam:

$$\sqrt{S_{X(2)}^m} = \frac{1.68}{\sigma^{\frac{1}{2}} \lambda^{\frac{1}{2}} B} \left( \frac{\rho}{E} \right)^{\frac{1}{8}} f_0^{-\frac{3}{4}} t^{-\frac{3}{4}} w^{-\frac{1}{2}} \left[ k_B T + \frac{S_V^a}{R_L} \left( \frac{1-0.75\alpha}{1-\alpha} \right) \right]^{\frac{1}{2}}$$

$$\sqrt{S_{X(2)}^m} = \frac{1.71}{\sigma^{\frac{1}{2}} \lambda^{\frac{1}{2}} B} \left( \frac{\rho}{E} \right)^{\frac{1}{2}} \left( \frac{L}{t} \right)^{\frac{3}{2}} w^{-\frac{1}{2}} \left[ k_B T + \frac{S_V^a}{R_L} \left( \frac{1-0.75\alpha}{1-\alpha} \right) \right]^{\frac{1}{2}} \quad \text{for } w$$

- 15 = t

Note that the sensitivity is independent of frequency for beams of constant aspect ratio.

In the preceding calculation we replaced the width  $w$  of the device for the width of the electrode. This presumes that there is only a single electrode, and that the device is driven by another means, or is used in a passive measurement. If magnetomotive drive and detect are used simultaneously, the calculation is identical in all respects except

5 that  $w$  must be replaced by  $w/3$ , the approximate width of an individual electrode.

### *Comparison to thermal noise*

The ultimate noise floor for a measurement of a mechanical transducer is its intrinsic thermal fluctuations. The spectral density of displacement noise corresponding

10 to thermal fluctuations of a mechanical transducer has a Lorentzian line shape, with a value on resonance given by:

$$\sqrt{S'_X} = \sqrt{\frac{4k_B T Q}{\xi k \omega_0}}$$

In the particular case of a straight doubly-clamped beam,

$$\sqrt{S'_X} = 0.200 \frac{(k_B T Q)^{\frac{1}{2}}}{E^{\frac{1}{8}} \rho^{\frac{3}{8}} t^{\frac{3}{4}} f_0^{\frac{5}{4}} w^{\frac{1}{2}}}$$

15 
$$\sqrt{S'_X} = 0.194 E^{-\frac{1}{2}} \left( \frac{L}{t} \right)^{\frac{3}{2}} w^{-\frac{1}{2}} f^{-\frac{1}{2}} \quad \text{for } w = t$$

For the two port measurement technique, then,

$$\sqrt{\frac{S_{X(2)}^M}{S'_X}} = \frac{3.15 \rho^{\frac{1}{2}}}{(\sigma \lambda Q)^{\frac{1}{2}} B} f_0^{\frac{1}{2}} \left[ 1 + \frac{S_V^a}{k_B T R_L} \left( \frac{1 - 0.75\alpha}{1 - \alpha} \right) \right]^{\frac{1}{2}}$$

From the above expression, the level of amplifier noise required to detect thermomechanical fluctuations decreases roughly as  $1/f_0$ , and in the limit of small  $\alpha$  is independent of other geometric factors. Neglecting the  $\alpha$  term, we can solve for the amplifier input noise necessary to achieve the thermomechanical limit:

$$S_V^a = 0.1 \frac{k_B TR_L \sigma \lambda Q B^2}{\rho f_0}$$

Although its overall sensitivity scales well to high frequencies, the frequency range of magnetomotive detection is fundamentally limited by the necessary measurement circuit. In the following section we will determine where this frequency limit lies for practical systems.

#### *Numerical example:*

Typical low-noise RF amplifiers with input impedance  $R_L = 50 \Omega$  have noise figures ranging from 0.3dB to 1.0dB for a source impedance of  $50\Omega$ . In the two-port detection circuit described in this report, the amplifier sees  $50\Omega$  through the impedance transformation, so the noise figure (NF.) can be converted to a power spectral density by the following equation:

$$S_V^a = (4k_B TR_L) 10^{NF/10dB}$$

This gives an effective noise voltage  $S_V^a$  across  $50\Omega$ , which includes both the voltage and current noise at the amplifier's input. For the quoted noise figures, the amplifier noise voltage ranges from  $0.93 \text{ nV} / \sqrt{\text{Hz}}$  to  $1.0 \text{ nV} / \sqrt{\text{Hz}}$ , assuming the amplifier is at room temperature. For a cryogenic amplifier at 4K, the noise level drops to  $.12 \text{ nV} / \sqrt{\text{Hz}}$ .



Consider a silicon beam of square cross-section with the following electrical parameters:  $\lambda = 0.1$ ,  $R_L = 50\Omega$ ,  $\sigma = 1.6 \times 10^7 / \Omega\text{-m}$ , in a magnetic field  $B = 8\text{T}$ . The two-port detection sensitivity is:

$$\sqrt{S_{X(2)}^m} = 1.72 \times 10^{-6} f_0^{-\frac{3}{4}} t^{-\frac{3}{4}} w^{-\frac{1}{2}} \left[ k_B T + \frac{S_V^a}{R_L} \left( \frac{1-0.75\alpha}{1-\alpha} \right) \right]^{\frac{1}{2}}$$

$$5 \quad \sqrt{S_{X(2)}^m} = 7.2 \times 10^{-11} f_0^{\frac{1}{2}} \left( \frac{L}{t} \right)^{\frac{5}{2}} \left[ k_B T + \frac{S_V^a}{R_L} \left( \frac{1-0.75\alpha}{1-\alpha} \right) \right]^{\frac{1}{2}}$$

where

$$\alpha = 850000 f_0^{\frac{1}{2}} t^{\frac{3}{2}}$$

and

$$\alpha = 7.0 \times 10^{11} \left( \frac{t}{L} \right)^3 f_0^{-1}$$

10 The thermomechanical noise is:

$$\sqrt{S_X^{th}} = 4.27 \times 10^{-4} (k_B T Q)^{\frac{1}{2}} f_0^{-\frac{5}{4}} t^{-\frac{5}{4}}$$

$$\sqrt{S_X^{th}} = 5.01 \times 10^{-9} (k_B T Q)^{\frac{1}{2}} \left( \frac{L}{t} \right)^{\frac{5}{2}}$$

Fig. 41 summarizes the sensitivity calculations in a graph of the sensitivity of the two-port magnetomotive detection technique as a function of frequency, compared to thermomechanical noise, for straight doubly-clamped silicon beams of  $Q = 10000$  and different thicknesses, measured in a magnetic field of 8 T. Note that the frequency at which thermomechanical noise can be measured is dependent only on parameters of

the electrical circuit.

Fig. 42 is a graph of the input noise level required of a  $50\Omega$  amplifier for magnetomotive sensitivity limited by thermomechanical noise, as a function of the conductivity of the electrode. The device is a straight doubly-clamped silicon beam with  $Q = 10000$  in a magnetic field of 8 T, and its electrode is  $1/10$  as thick as the structure. Based on the expression derived earlier for the amplifier input noise, the best way to extend the magnetomotive technique into the GHz frequency range would be to increase the conductivity of the detection electrode. The plot in Fig. 42 shows the effectiveness of this approach.

The magnetomotive technique is a very powerful tool for the detection of nanomechanical transducers in motion. It attains high sensitivity at frequencies up to and over 1 GHz, and has a large linear dynamic range. The physical principles underlying its effectiveness are very basic, enabling straightforward analysis of measured signals. With a simple readout circuit and a standard RF amplifier, magnetomotive detection can attain the sensitivity limit of thermomechanical fluctuations for a nanomechanical transducer operating at 1 GHz.

#### *NEMS Fabrication Using Si and GaAs Membranes*

Si and GaAs membranes can be fabricated using bulk micromachining techniques. In both cases, backside-processing using anisotropic selective etchants produces a suspended membrane of various widths and dimensions which can be further micromachined into a wide array of devices. While the basic method for each process is the same, the different crystallographic natures of the two materials require two distinct procedures.

### *Si Membrane Fabrication*

Stiction in microelectromechanical systems (MEMS) has been a major failure mode ever since the advent of surface micromachining in the 80s of the last century due to large surface-area-to-volume ratio. When the devices are scaled down to

5    nanoelectromechanical systems (NEMS), the stiction poses an even more challenging issue during in the fabrication process. By patterning the NEMS devices in predefined membrane, suspended nano-structures are no longer in close proximity to the substrate. Release related stiction is effectively prevented during drying. Higher yield of NEMS devices are therefore achieved.

10        The membrane substrates are also beneficial for high resolution lithography since backscattering in the substrate during the exposure of the pattern is much reduced. We have proved that nanometer scale pattern can be easily defined through electron beam lithography.

      The procedure for processing Si membranes is outlined in Fig. 43a- 43d.

15    Membrane fabrication begins with a material comprised of a silicon epilayer 104 and a .4 $\mu$ m thick implanted SiO<sub>2</sub> layer 106, bonded to a Si substrate 108 as depicted in Fig. 43a. A highly anisotropic *KOH* wet etch is used to remove a region 110 of the bulk Si substrate 108 from the backside of a sample. The selective etch characteristics of *KOH* allow the SiO<sub>2</sub> to serve as an etch stop layer, which ensures a smooth backside and a

20    well-defined and uniform membrane thickness.

### *Etch Anisotropy of Silicon*

*KOH* etches Si precisely along its crystal planes, forming a pyramidal etch window 110 which forms a sidewall angle of 125° as depicted in Fig 43b. Undercutting of the

25    mask is negligible for our purposes. This precise anisotropy allows membranes of any

size to be constructed fairly easily. The mask is comprised of a series of squares of the appropriate sizes, separated by lines along the cleave planes to facilitate multiple sample processing and easy cleaving into individual dies once the process is finished.

## 5 *Membrane Fabrication*

Due to the aggressive nature of the KOH etch, low stress (Si-rich)  $\text{Si}_3\text{N}_4$  is used as a mask. Both sides of the wafer are coated with 600Å of  $\text{Si}_3\text{N}_4$  via low pressure chemical vapor deposition (LPCVD), creating a pinhole free protection layer 112 for the Si epilayer 104, as well as a masking layer 114 for the backside. The mask 112 is defined in the nitride by photolithography and subsequent etching in an electron cyclotron resonance (ECR) system, using a mixture of 10 standard cubic centimeters per minute (sccm) of Ar and 20 sccm  $\text{NF}_3$  for 2 minutes. A layer of photoresist layer (not shown) should be spun on the epilayer side for protection before etching to ensure that the silicon nitride coating 112 is not damaged.

The bulk Si etch is performed in a 30% KOH solution, held at 80°C and mixed just prior to etching. This volume ratio yields a maximum etch rate of approximately 1.4  $\mu\text{m}/\text{min}$ , requiring an etch time of just over 6 hours before the  $\text{SiO}_2$  layer 106 is reached. KOH etches  $\text{SiO}_2$  at  $\sim 8 \text{ Å}/\text{min}$ , leaving ample time to stop the etch before doing any damage to the Si epilayer 104.

The  $\text{SiO}_2$  sacrificial layer 106 is removed in a 10% HF solution, with an etch rate of  $\sim 340 \text{ Å}/\text{min}$  as depicted in Fig. 43c. Undercutting of the  $\text{SiO}_2$  layer 106 widens the membrane size no more than 4 $\mu\text{m}$  in both directions. Dilute HF etches  $\text{Si}_3\text{N}_4$  at a rate of  $\sim 3 \text{ Å}/\text{min}$ , removing only  $\sim 38 \text{ Å}$  of the mask 112 during the etch time of 12min.

The remainder of the  $\text{Si}_3\text{N}_4$  layer 112 is then removed in an 85%  $\text{H}_3\text{PO}_4$  bath kept at 160°C, for 6 minutes as depicted in Fig. 43d. The etch rate of  $\text{SiO}_2$  and Si in  $\text{H}_3\text{PO}_4$  is

negligible for our purposes, although some damage was observed on the Si layer 104 with etch times greater than 30 minutes.

It is possible that a small percentage of metal impurities in the solution can deposit on the underlying bare Si surface 104 through an electrochemical displacement plating reaction during the etch process. This is avoided by adding 5% HCl by weight to the solution to act as a chelating agent, leaving the etch characteristics unaffected. It should also be noted that as the solution evaporates, the etch rate slows considerably. For this reason the process should be carried out as soon as possible after the proper temperature is reached to ensure consistent results.

10

#### *GaAs Membrane Fabrication*

The procedure developed to create GaAs membranes is depicted in the side cross-sectional views of Figs. 44a – 44d. Processing begins with a material consisting of a bulk GaAs substrate 116, topped with a three electron beam epitaxial (MBE) grown layers: a 600 nm GaAs buffer layer 118, a 1  $\mu\text{m}$   $\text{Al}_{0.8}\text{Ga}_{0.2}\text{As}$  etch stop layer 120, and the appropriate GaAs epilayer 122 required for the desired final membrane thickness as depicted in Fig. 44a. Two anisotropic selective etches were investigated: a  $\text{NH}_4\text{OH}:\text{H}_2\text{O}_2$  solution, and a citric acid: $\text{H}_2\text{O}_2$  solution. Each etchant has its own characteristic etch profile and the advantages of each vary accordingly.

20

#### *Etch Anisotropy*

Anisotropic etching of GaAs presents some complications as compared to the previously described process in silicon with etch profiles differing along the two major crystal planes as well as with the etchant used. The  $\text{NH}_4\text{OH}$  solution produces well-

defined and smooth surfaces along the etched walls and floor as shown in the microphotograph of Fig. 45a, while the citric acid etches less uniformly on all surfaces as shown in the microphotograph of Fig. 46a. The undercut ratios for both etchants limit how small the final membrane dimensions can be, requiring a thinner substrate than that commercially provided in order to produce a membrane of reasonable size. The undercut ratio is defined as the ratio of the lateral etch rate to the vertical etch rate. The substrate 116 can be thinned down to 100 $\mu$ m, below which makes the sample very fragile and prone to breaking or chipping and less likely to survive later processing steps. Because the front side of the membrane is protected as described in processing steps below, it is possible to avoid the fragility problems due to substrate thickness by fabricating the desired device on the front surface before thinning the substrate and processing the membrane. This requires an infrared mask aligner to align the device with the membrane pattern before etching.

The etch rate of the  $\text{NH}_4\text{OH}:\text{H}_2\text{O}_2$  solution varies along different crystal planes depending on the volume ratios of etch products. The 1:30 solution, chosen for maximum selectivity, produces an obtuse sidewall angle of  $\sim 130^\circ$  in the  $(0\bar{1}0)$  plane and an acute sidewall angle of  $\sim 60^\circ$  in the  $(011)$  plane as depicted in Fig. 44b. In addition, significant undercutting also occurs, with an undercut ratio averaging  $\sim 0.5$  for both crystal planes. This serves to widen the mask window dimensions by  $\sim 1$  square micron for every micron of etched depth. The combination of the above characteristics constrains the dimension along the  $(011)$  plane to a minimum of  $\sim 200\mu\text{m}$  for an initial substrate thickness of 100 $\mu\text{m}$ .

The anisotropic etch characteristics of citric acid on GaAs differ somewhat from that of  $\text{NH}_4\text{OH}$ . For a volume ratio of 3:1 it also produces a sidewall angle of  $\sim 130^\circ$  in the  $[1\bar{1}0]$  direction, but an effective angle of  $90^\circ$  in the  $[011]$  direction as shown in the

microphotograph of Fig. 46b. The undercut ratios for the  $[0\bar{1}0]$  and  $[011]$  planes are  $1.2\mu\text{m}$  and  $1.5\mu\text{m}$  respectively. The combination of these two characteristics reduces the dimension constraint in the  $[011]$  direction to about  $150\mu\text{m}$  for an initial substrate thickness of  $100\mu\text{m}$ .

- 5        In cases where later device constraints require smaller membrane dimensions, the citric acid solution could be preferable to the  $\text{NH}_4\text{OH}$  solution. However, at the present conditions the etch rate approaches zero past a depth of about  $100\mu\text{m}$ . This requires the substrate to be thinned as much as possible creating a fragile sample which can be difficult to handle. Because the  $\text{NH}_4\text{OH}$  etchant can etch uniformly through thicknesses
- 10    of greater than  $600\mu\text{m}$  with well-defined and reproducible membrane dimensions, at the present time this solution is preferred when larger membrane sizes can be tolerated. Further experimentation with citric acid volume ratios and temperature conditions may prove the solution more useful at a later time.

## 15    *Membrane Fabrication*

### *Substrate thinning*

- The sample preparation process for both etch methods is identical. Membrane fabrication begins with thinning the GaAs substrate 116 to a thickness between  $300$  and  $100\mu\text{m}$  using a fast isotropic  $\text{H}_2\text{SO}_4:\text{H}_2\text{O}_2:\text{H}_2\text{O}$  wet etch in the volume ratio  $1:8:1$ . This
- 20    etches at approximately  $5\mu\text{m}/\text{min}$  and produces a reasonably smooth and sufficiently homogenous backside surface for our purpose. A piece of the material a few millimeters on a side is prepared, which will later be cleaved into smaller samples for individual membrane processing.

- A layer of photoresist 124 is spun on the front side to protect the epilayer 120
- 25    before waxing the material face down to a glass coverslip. AZ 4330 photoresist is used,

and care should be taken not to heat the sample and wax above 130°C as it makes the photoresist extremely difficult to remove later in the process. Once the wax has hardened, a small cotton swab with acetone can be used to gently remove the photoresist residue from the backside of substrate 116.

It should be noted that the etch rate is extremely sensitive to temperature. As some heating occurs when the etchant components are mixed, the solution is left for an hour to return to room temperature before immersing the sample. Also due to this temperature sensitivity, normal room temperature fluctuations can result in a somewhat unstable etch rate, varying by as much as 20%. Because removing the sample from the solution periodically to determine thickness can produce markedly different etch times subsequent etch rates, a vertical micrometer is helpful in achieving the exact desired material thickness. Once this thickness is reached, the sample is rinsed thoroughly in DI water and left in acetone to dissolve the wax.

## 15 Etch methods

Once the wax is removed, photoresist 126 is again spun on the front for protection. The backside is then flood exposed in a mask aligner and developed to remove residual resist. AZ 4330 photoresist 126 is spun on the backside of the sample at 2750 rpm and baked for 1 min at 95°C, producing a resist layer about 5 μm thick. The etch mask corresponding to the final membrane dimensions is then defined relative to the proper crystal planes. After the pattern is developed it is post-baked at 115°C for 2min, while waxing the sample epilayer side down to a glass microscope slide.

The  $\text{NH}_4\text{OH}$  solution used is comprised of  $\text{NH}_4\text{OH}$  and  $\text{H}_2\text{O}_2$  in the volume ratio of 1:30 for greatest selectivity ( $\sim 100$ ), and is freshly mixed prior to etching. The reaction is

25 diffusion-rate limited and spraying it onto the sample serves to circulate and mix the



solution, as well as mechanically remove etch products. It should be noted that the use of a Teflon sample holder is important to ensure the greatest selectivity. When the AlGaAs sacrificial layer is reached, the etched window becomes transparent through the top two layers and orange in color. The etch is allowed to continue for ~30 seconds to assure complete removal of the underlying GaAs layer, and the sample is rinsed thoroughly in DI water to ensure removal of all etch products.

The citric acid solution previously mentioned can also be used to remove the bulk substrate. This is reaction-rate limited, and therefore used as a simple bath. Citric acid monohydrate is mixed 1:1 with DI water by weight one day in advance to ensure complete dissolution. This solution is then mixed in a 3:1 volume ratio with  $H_2O_2$ , and allowed to rest for approximately 20 minutes to return to room temperature. The sample is immersed in the bath until the transparent window is seen (just over 6 hours for an initial substrate thinned to  $100\mu m$ ), and rinsed thoroughly.

At this point the AlGaAs layer is removed by immersing the sample, still attached to the glass slide, in 20%HF for 1min 15s, with a selectivity of AlGaAs to GaAs of greater than 107 as depicted in Fig. 44c. When the AlGaAs layer has been completely removed, a faint ring can be optically seen around the membrane, indicating undercutting of the sacrificial layer. To complete the process, the sample is left in acetone overnight to dissolve the wax, transferred to isopropyl alcohol, and gently blown dry resulting in the structure of Fig. 44d.

A process has been developed to produce membrane structures out of silicon and gallium arsenide using bulk micromachining methods. Both processes utilize selective etching anisotropic etching systems. For the Si system, a well-defined KOH etchant was characterized, which is selective to Si over  $SiO_2$ . For the GaAs system,  $NH_4OH$  and citric acid solutions were characterized, both of which are selective to GaAs over

AlGaAs. It was found that the preferred etchant for both reproducibility and durability is  $\text{NH}_4\text{OH}$ , unless future device constraints require membrane dimensions less than  $150\mu\text{m}$ .

## 5 *NEMS Array Scalar Analyzers/ Correlators*

Fig. 47 illustrates basic concepts behind a NEMS array spectrum analyzer 128. In this conceptualization the analog of "resonant reeds" are piezoresistive NEMS cantilevers, as pictured in Fig. 47. The elements 130 forming the array 128 have lengths that are staggered (here denoted as  $L_i, \dots, L_k$ ), thus yielding overall resonant  
 10 response that covers some desired, preprogrammed spectral range. Here each element 130 is pictured as being separately driven and sensed, however all share a common ground electrode 132. It is noteworthy that even simpler readouts are possible. The signal is pictured here as being delivered from a common transmission line 134 with local stubs 136 to provide electrostatic actuation at each element 130.  
 15 Note that a difference in thickness difference between the drive electrode 138 and the cantilever tips 140 in Fig. 47 provides requisite the out-of-plane electric fields for inducing mechanical motion in this direction.

Fig. 47 represents a realization where individual, uncoupled elements provide the functionality. It is also possible to have collective mechanical modes in a coupled  
 20 array of mechanical elements. This provides for a broad class of optoelectromechanical array spectrum analyzers. One simple realization from this family is conceptually depicted in Fig. 48 where a plurality of interdigitated or otherwise arrayed and interacting beams or cantilevers 210 as shown in Fig. 48a are disposed between two opposing T-frames 212. Here the Fourier components present in the electrical signal  
 25 waveform, denoted as  $v(t)$ , parametrically drive the collective modes of the array. This

motion, in turn, modulates the strength of the diffracted orders of light from a laser 214 coupled to device 10 by means of an optic fiber 216 collimated by collimator 218 and transmitted through the array 128 which is, in essence, a time-varying optical diffraction grating. These orders can be read out continuously and therefore can provide real-time  
5 spectral analysis of the electrical waveform,  $v(t)$  at input 220.

### *NEMS Array Chemical/Biological Sensors*

Two groups have pioneered *MEMS* based electromechanical "nose" devices. The efforts have primarily been directed toward the sensing of gaseous analytes and  
10 fluidic analytes. There are two modes of operation that stems from two distinct physical mechanisms of interaction. The first mode, which is the basis for the recent work from both groups, is based upon the induction of differential strain in the cantilevers from an overlayer that swells or shrinks upon exposure to the analytes. If this overlayer coats only one face of the cantilever, the swelling or shrinkage of the overlayer results in  
15 bending, which is then detected optically.

The second mode of sensing is based upon mass loading, and the resultant change in the total inertial mass of the sensor, which can be detected as a resonant frequency shift.

There are significant and compelling reasons for scaling these ideas down into  
20 the realm of NEMS arrays. Most significant is that the sensitivity of a "electromechanical nose" can be greatly enhanced due to the smaller mass of NEMS elements, and also by the further improvements that can be derived from the enhanced strain sensitivity, mass sensitivity, compliance, and operating frequency of nanoscale mechanical elements. A concrete example of this is given by our recent work on ultrasensitive

NEMS mass sensing.

Fig. 49 illustrates a conceptualization of a NEMS array electronic nose. Each element 142 within an array of separately transduced piezoresistive cantilevers 144 is surface loaded with a film providing sensitivity to a particular target analyte. In this conceptualization, adjacent electrostatic drive electrodes 146 allow separate excitation of the chemically functionalized elements. This would require individual connections to each drive electrode.

Another means for addressing each element 142 is shown in Fig. 50; this employs a single transmission line 130 and a swept signal yielding addressability in the frequency domain if the cantilevers 144 are designed with staggered lengths as depicted in Fig. 50.

#### *NEMS Array Infrared Detectors/Imagers*

In the illustrated embodiment IR imagers are based upon NEMS arrays 128. The significant reductions in size will provide immense pay-offs in terms of sensitivity and response time. One possible device layout is shown in Fig. 50. Here the resonant frequencies of the individual elements are staggered by means of lithographically patterned variations in the lengths of the IR absorber (?). AC readout of the strain-induced bending arising from IR absorption in the "absorbers" is detected as a frequency shift. This shift is the direct consequence of a resonant frequency for each element that is dependent upon its average position. This position dependence arises from a static DC voltage applied to each element's drive electrode in addition to the RF drive signal itself. This DC voltage bias translates into an electrostatic term in each cantilever's potential energy, resulting in a position dependent resonant frequency. In this particular conceptualization, we also envisage fast interrogation of the large

number of array elements by stepped frequency excitation of the individual resonant elements. This allows individual addressability via a single transmission line. It is quite reasonable to envisage frequency-multiplexing the readouts in similar manner, by AC coupling the piezoresistors to a common readout transmission line.

5           Many alterations and modifications may be made by those having ordinary skill in the art without departing from the spirit and scope of the invention. Therefore, it must be understood that the illustrated embodiment has been set forth only for the purposes of example and that it should not be taken as limiting the invention as defined by the following claims. For example, notwithstanding the fact that the elements of a claim are  
10   set forth below in a certain combination, it must be expressly understood that the invention includes other combinations of fewer, more or different elements, which are disclosed in above even when not initially claimed in such combinations.

          The words used in this specification to describe the invention and its various embodiments are to be understood not only in the sense of their commonly  
15   defined meanings, but to include by special definition in this specification structure, material or acts beyond the scope of the commonly defined meanings. Thus if an element can be understood in the context of this specification as including more than one meaning, then its use in a claim must be understood as being generic to all possible meanings supported by the specification and by the word itself.

20           The definitions of the words or elements of the following claims are, therefore, defined in this specification to include not only the combination of elements which are literally set forth, but all equivalent structure, material or acts for performing substantially the same function in substantially the same way to obtain substantially the same result. In this sense it is therefore contemplated that an equivalent substitution of  
25   two or more elements may be made for any one of the elements in the claims below or

that a single element may be substituted for two or more elements in a claim. Although elements may be described above as acting in certain combinations and even initially claimed as such, it is to be expressly understood that one or more elements from a claimed combination can in some cases be excised from the combination and that the  
5 claimed combination may be directed to a subcombination or variation of a subcombination.

Insubstantial changes from the claimed subject matter as viewed by a person with ordinary skill in the art, now known or later devised, are expressly contemplated as being equivalently within the scope of the claims. Therefore, obvious  
10 substitutions now or later known to one with ordinary skill in the art are defined to be within the scope of the defined elements.

The claims are thus to be understood to include what is specifically illustrated and described above, what is conceptionally equivalent, what can be obviously substituted and also what essentially incorporates the essential idea of the  
15 invention.